

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 6016.11 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataacatataattcatc.....cgttgtgctgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgtg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2032	100.0	2032	6	AR257469	AR257469 Sequence
2	2032	100.0	2032	6	AX148742	AX148742 Sequence
3	2032	100.0	2032	8	ARQARAH	I34402 Arachis hyp
4	2032	100.0	2041	6	AR257470	AR257470 Sequence
5	2019.2	99.4	2032	6	AX155331	AX155331 Sequence
6	1930	95.0	1930	6	BD172109	BD172109 Peanut al
7	1930	95.0	1930	6	AR257463	AR257463 Sequence
8	1930	95.0	1930	6	BD107898	BD107898 Methods a
9	1805.2	88.8	1978	6	AX155332	AX155332 Sequence
10	1752.4	86.2	1949	6	AR257467	AR257467 Sequence
11	1752.4	86.2	1949	8	ARQARAH	I38853 Arachis hyp
12	1608	79.1	2674	8	AF432231	AF432231 Arachis h
13	1395.4	68.7	2332	8	AY581852	AY581852 Arachis h
14	1375.4	67.7	1418	8	AY581851	AY581851 Arachis h
15	954.4	47.0	1035	8	AY581850	AY581850 Arachis h
16	891.2	43.9	1047	8	AY581849	AY581849 Arachis h
17	750	36.9	750	6	AR257466	AR257466 Sequence
18	468.6	23.1	1257	8	LCU551424	AJ551424 Lens culi
19	465.8	22.9	1433	8	PSVIC	Y00722 Pisum sativ

20	461	22.7	1561	8	VFVIC	Y00462 Vicia faba
21	455.4	22.4	1248	8	LCU551425	AJ551425 Lens culi
22	449	22.1	1248	8	AJ626897	AJ626897 Pisum sat
23	449	22.1	1596	8	PSVICK	X67429 Pisum sativ
24	446	21.9	1791	8	AY500372	AY500372 Lupinus a
25	444.2	21.9	1248	8	AJ626898	AJ626898 Pisum sat
26	435.4	21.4	1548	8	VNVICLN	Z71987 V.narbonens
27	408.4	20.1	1251	8	AB008679	AB008679 Glycine m
28	406.8	20.0	1320	6	BD175903	BD175903 Regulatio
29	406.8	20.0	1320	6	AR202568	AR202568 Sequence
30	406.8	20.0	1320	6	AR486455	AR486455 Sequence
31	404	19.9	1632	8	AB008678	AB008678 Glycine m
32	404	19.9	1818	6	BD175901	BD175901 Regulatio
33	404	19.9	1818	6	AR202566	AR202566 Sequence
34	404	19.9	1818	6	AR486453	AR486453 Sequence
35	403.6	19.9	1955	8	GMASBC	X17698 Soybean mRN
36	403.6	19.9	1477	8	AB030840	AB030840 Glycine m
37	403.6	19.9	1477	8	AB030841	AB030841 Glycine m
38	402	19.8	1680	8	AB008680	AB008680 Glycine m
39	402	19.8	1884	8	AB113351	AB113351 Glycine m
40	400.8	19.7	2070	8	AB030839	AB030839 Glycine m
41	398.8	19.6	2042	8	AB030838	AB030838 Glycine m
42	384	18.9	1336	8	AY221105	AY221105 Glycine m
43	378.8	18.6	2069	8	PSA276875	AJ276875 Pisum sat
44	347.8	17.1	1858	8	VNCONVN	Z71986 V.narbonens
45	339.4	16.7	1494	8	GMU14006	U14006 Glycine max

ALIGNMENTS

RESULT 1
LOCUS AR257469 2032 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 21 from patent US 6486311.
ACCESSION AR257469
VERSION AR257469.1 GI:27307482
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2032)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannan,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 21 26-NOV-2002;
FEATURES Location/Qualifiers
source 1..2032
/organism="unknown"
/mol_type="genomic DNA"

Query Match	100.0%	Score 2032;	DB 6;	Length 2032;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGGAGGAGCAATGAGAGGGAG	60	
DB	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGGAGGAGCAATGAGAGGGAG	60	
QY	61	GGTTTCTCCAGTCGCTGTGCTAGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120	
DB	61	GGTTTCTCCAGTCGCTGTGCTAGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCA	120	
QY	121	TGCCAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCGAGAGGTGCCTCCA	180	
DB	121	TGCCAAGTCATCACCTTACAGAGAAACAGAGAACCCCTCGCCGAGAGGTGCCTCCA	180	
QY	181	GAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTCTCGCTGCACCAA	240	
DB	181	GAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTCTCGCTGCACCAA	240	
QY	241	GCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGGCACCAACCA	300	

[illegible]

Db	1321	CCAGCTTGAAGCACTGGACATGATGCTCACTCGTGTAGAGATCAAAGAAAGGAGCTTTGAT	1380
Qy	1381	GCTCCCACTTCAAATCAAAGGCCATGGTTATCGTCGTCAAAAGAGAACTGGAAA	1440
Db	1381	GCTCCCACTTCAAATCAAAGGCCATGGTTATCGTCGTCAAAAGAGAACTGGAAA	1440
Qy	1441	CCTTTGAATCTGTGGCTGTAAAGAAAGCAACACAGAGGGGACGGCGGGAAGAGGAA	1500
Db	1441	CCTTTGAATCTGTGGCTGTAAAGAAAGCAACACAGAGGGGACGGCGGGAAGAGGAA	1500
Qy	1501	GGACGAAGAACGAAGAGGAGGGAAGTAACAGAGAGGTGCTGTAGGTACACAGCAGAGTT	1560
Db	1501	GGACGAAGAACGAAGAGGAGGGAAGTAACAGAGAGGTGCTGTAGGTACACAGCAGAGTT	1560
Qy	1561	GAAGGAAGGCGATGTGTTCTATCATGTCAGCAGCTCATCCAGTAGCCCATCAACGTTCTCTC	1620
Db	1561	GAAGGAAGGCGATGTGTTCTATCATGTCAGCAGCTCATCCAGTAGCCCATCAACGTTCTCTC	1620
Qy	1621	CGAACTCCATCTGCTTGGCTTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTTC	1680
Db	1621	CGAACTCCATCTGCTTGGCTTTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTTC	1680
Qy	1681	AGGTGATAAGGACAAATGTGATAGACAGATAGAGAAGCAAGCGAAGGATTTAGCATTTCCC	1740
Db	1681	AGGTGATAAGGACAAATGTGATAGACAGATAGAGAAGCAAGCGAAGGATTTAGCATTTCCC	1740
Qy	1741	TGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAACACAGAGGAATCTCAGCTTTGTGAG	1800
Db	1741	TGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAACACAGAGGAATCTCAGCTTTGTGAG	1800
Qy	1801	TGCTCGTCCTCAATCTCAATCTCAATCTCGTCTCCTGAGAAAGAGTCTCCTCAGAA	1860
Db	1801	TGCTCGTCCTCAATCTCAATCTCAATCTCGTCTCCTGAGAAAGAGTCTCCTCAGAA	1860
Qy	1861	AGAGGATCAAGAGAGGAGAAAAACCAAGAGGGAAGGGTCCACTCCTTTCAATTTTGAGGC	1920
Db	1861	AGAGGATCAAGAGAGGAGAAAAACCAAGAGGGAAGGGTCCACTCCTTTCAATTTTGAGGC	1920
Qy	1921	TTTTTAACCTGAGAAATGGAGGCAACTGTTTATGATCGATAATAGATACAGCTTTTGTACT	1980
Db	1921	TTTTTAACCTGAGAAATGGAGGCAACTGTTTATGATCGATAATAGATACAGCTTTTGTACT	1980
Qy	1981	CTACTATCCAAAACTTATCATCAATAAATAAAAAAGCTTTGTGCGTGTGTTCTCC	2032
Db	1981	CTACTATCCAAAACTTATCAATAAATAAAAAAGCTTTGTGCGTGTGTTCTCC	2032

		Query Match	100.0%;	Score 2032;	DB 6;	Length 2032;		
		Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	0;	Indels	0;
		Matches 2032;	Conservative	0;	0;	0;	0;	0;
Qy	1	AATAATCATATATATATCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGGGAG	60					
Db	1	AATAATCATATATATATCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGGGAG	60					
Qy	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTCGCTTCAGTTTCTGGAACGCA	120					
Db	61	GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTCGCTTCAGTTTCTGCAACGCA	120					
Qy	121	TGCCAAGTCATCACTTTACAGAGAAACACAGAGAACCCCTGGCCACAGAGGTGCCTCCA	180					
Db	121	TGCCAAGTCATCACTTTACAGAGAAACACAGAGAACCCCTGGCCACAGAGGTGCCTCCA	180					
Qy	181	GAGTTGTCAAACAGAAACCGGATGATCTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCA	240					
Db	181	GAGTTGTCAAACAGAAACCGGATGATCTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCA	240					
Qy	241	GCTCGAGTATGATCCTGTTGTGTCTATGATCTCTGAGGACACACTGGGACCAACCA	300					
Db	241	GCTCGAGTATGATCCTGTTGTGTCTATGATCTCTGAGGACACACTGGGACCAACCA	300					
Qy	301	ACGTTCCCTCCAGGGAGCGGACACGCTGGCCGCAACCCGAGACTACGATGATGACCG	360					
Db	301	ACGTTCCCTCCAGGGAGCGGACACGCTGGCCGCAACCCGAGACTACGATGATGACCG	360					
Qy	361	CCGTCAAACCCGAGAGAGGAGCGCCGATGGGACACAGCTGGACCGAGGGAGCGTGA	420					
Db	361	CCGTCAAACCCGAGAGAGGAGCGCCGATGGGACACAGCTGGACCGAGGGAGCGTGA	420					
Qy	421	AAGAGAAGACTGAGAGACAAACAGAGAAGATTGGAGCGGACCAAGTCAATCAGAGCC	480					
Db	421	AAGAGAAGACTGAGAGACAAACAGAGAAGATTGGAGCGGACCAAGTCAATCAGAGCC	480					
Qy	481	ACGGAATAAAGCCGCAAGAGAGAGAGAGAACAGAGTGCGGACACACAGTAGCCA	540					
Db	481	ACGGAATAAAGCCGCAAGAGAGAGAGAGAACAGAGTGCGGACACACAGTAGCCA	540					
Qy	541	TGTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC	600					
Db	541	TGTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC	600					
Qy	601	CCGCTACGGGAACCAAAACCGTAGGATCCGGGCTCTGCAGAGGTTTGAACAAAGGTCAAG	660					
Db	601	CCGCTACGGGAACCAAAACCGTAGGATCCGGGCTCTGCAGAGGTTTGAACAAAGGTCAAG	660					
Qy	661	GCAAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCGAGCCAAACCTTAACACTCT	720					
Db	661	GCAAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCGAGCCAAACCTTAACACTCT	720					
Qy	721	TGTTCTTCCAGCAGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCCAC	780					
Db	721	TGTTCTTCCAGCAGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCCAC	780					
Qy	781	CGTGACGCTAGCAATAGGCAATACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACT	840					
Db	781	CGTGACGCTAGCAATAGGCAATACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACT	840					
Qy	841	CAGAAATCCCATCCGTTTCAATTTCTTACATCTTGAACCGCCATGACAAACAGAACTCTAG	900					
Db	841	CAGAAATCCCATCCGTTTCAATTTCTTACATCTTGAACCGCCATGACAAACAGAACTCTAG	900					
Qy	901	AGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTTCCTCCCGG	960					
Db	901	AGTAGCTAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTTCCTCCCGG	960					
Qy	961	GAGCAGCCGAGACCAATCATCTTTCAGAGGCTTCAGCAGGAATACTGTTGAGGCCCG	1020					
Db	961	GAGCAGCCGAGACCAATCATCTTTCAGAGGCTTCAGCAGGAATACTGTTGAGGCCCG	1020					
Qy	1021	CTTCAATGCGGAATTCATAGATACGGAGGCTGCTGTTAGAGAGATGACGAGGCTGA	1080					

1021	CTTCAATGCGGAATTCATAGATACGGAGGCTGCTGTTAGAGAGATGACGAGGCTGA	1080
1081	GCAAGAGGAGAGAGGCGCAGAGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT	1140
1081	GCAAGAGGAGAGAGGCGCAGAGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT	1140
1141	GATAGTCAAAAGTGTCAAAAGGACAGCTTGAAGAACTTACTTAAGACAGCTAAATCCGCTCTC	1200
1141	GATAGTCAAAAGTGTCAAAAGGACAGCTTGAAGAACTTACTTAAGACAGCTAAATCCGCTCTC	1200
1201	AAAGAAAGGCTCCGAAGAGAGGAGATATCACAAACCAATCAACTTTGAGAGAGGCGA	1260
1201	AAAGAAAGGCTCCGAAGAGAGGAGATATCACAAACCAATCAACTTTGAGAGAGGCGA	1260
1261	GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCC	1320
1261	GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAAGAACCC	1320
1321	CCAGCTTCAGGACCTGGACATGATGCTCACTGCTAGAGATCAAAAGAGGAGCTTTGAT	1380
1321	CCAGCTTCAGGACCTGGACATGATGCTCACTGCTAGAGATCAAAAGAGGAGCTTTGAT	1380
1381	GCTCCCACTCTTCAACTCAAAAGGCCATGTTATCGTCTGTCGTCACAAAGAGAACTGAAA	1440
1381	GCTCCCACTCTTCAACTCAAAAGGCCATGTTATCGTCTGTCGTCACAAAGAGAACTGAAA	1440
1441	CCTTGAACCTCGTGGCTGTAAGAAAAGAGCAACAAAGAGGGGACGGCGGGAAGAGGA	1500
1441	CCTTGAACCTCGTGGCTGTAAGAAAAGAGCAACAAAGAGGGGACGGCGGGAAGAGGA	1500
1501	GGACGAGACGAAAGAGAGAGGAAAGTAACAGAGAGGTCGCTAGGTACACAGCGAGGTT	1560
1501	GGACGAGACGAAAGAGAGAGGAAAGTAACAGAGAGGTCGCTAGGTACACAGCGAGGTT	1560
1561	GAAAGAGCGGATGCTTCAATCATGCGAGCAGCTCATCAGTAGCCATCAAGCTTCCCTC	1620
1561	GAAAGAGCGGATGCTTCAATCATGCGAGCAGCTCATCAGTAGCCATCAAGCTTCCCTC	1620
1621	CGAATCTCAATCTCTGCTTGGCTTCGGTATCAACGCTGAAAACCAACAGAAATCTTCTTGC	1680
1621	CGAATCTCAATCTCTGCTTGGCTTCGGTATCAACGCTGAAAACCAACAGAAATCTTCTTGC	1680
1681	AGGTGATAGGACAAATGTGATAGACAGATAGAGAGGAGGAGGAGGATTTAGCATCCC	1740
1681	AGGTGATAGGACAAATGTGATAGACAGATAGAGAGGAGGAGGAGGATTTAGCATCCC	1740
1741	TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGGAAATCTCAGTTGTAG	1800
1741	TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGGAAATCTCAGTTGTAG	1800
1801	TGCTCGTCTCAATCTCAATCTCCGTCGTCCTCGTCTCTGAGAAAGAGTCTCCTGAGAA	1860
1801	TGCTCGTCTCAATCTCAATCTCCGTCGTCCTCGTCTCTGAGAAAGAGTCTCCTGAGAA	1860
1861	AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1920
1861	AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1920
1921	TTTAACTCAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1980
1921	TTTAACTCAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1980
1981	CTACTATCCAAAACCTTATCAATAAATAAAGGTTTGTGCGTTGTTCTCC	2032
1981	CTACTATCCAAAACCTTATCAATAAATAAAGGTTTGTGCGTTGTTCTCC	2032

RESULT 3
 ARQARAH
 LOCUS
 DEFINITION
 ACCESSION

2032 bp mRNA linear PLN 24-MAY-1996
 Arachis hypogaea (clone P41b) Ara h I mRNA, complete cds.

L34402

VERSION	L34402.1	GI:602435	
KEYWORDS	allergen.		
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
REFERENCE	1 (bases 1 to 2032)		
AUTHORS	Burks, A.W., Cockrell, G., Stanley, J.S., Helm, R.M. and Bannion, G.A.		
TITLE	Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity		
JOURNAL	J. Clin. Invest. 96 (4), 1715-1721 (1995)		
MEDLINE	96013631		
PUBMED	7560062		
COMMENT	On Dec 16, 1994 this sequence version replaced gi:508640. Original source text: Arachis hypogaea (strain Florunner).		
FEATURES	Location/Qualifiers		
source	1..2032		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/strain="Florunner"		
	/db_xref="taxon:3818"		
	/clone="F41b"		
	/dev_stage="seed"		
gene	1..2032		
	/gene="Ara h I"		
5'UTR	<1..49		
	/gene="Ara h I"		
CDS	50..1930		
	/gene="Ara h I"		
	/codon_start=1		
	/protein_id="AAB00861.1"		
	/db_xref="GI:602436"		
	/translation="MRGVSPLEMLLILVLASVSATHAKSSPYOKKTENPCQRLQ SCQEPDDLKQKACSRCTKLEYDFRCVDPDGHGTGTTNQRPSPGRTNGRPQGYDD DRQPRREGGWSGAPREREDWRPDRPSPHQPKRIPBREGEGEWGT PSHVREETSRRNPFYFSPRSTRYGNQGRIRVLQRFDQSRQFQNLQNRHIVQIE AKPNTLVLPHKADANILVIQOQATVTVANGNRKNSFNLDRGHALRIPSGFISVLN RDHONLRVAKLSMPNTPGQFDPFPASSRQSSVLQGFSRNTLEAFAFNFRIR VLLENAGGEQERQORWSTRSSENNEGVIVKVSKEHVVELTKHAKSVSKGSBEEG DITPNLNRGEGPDSUNFGKLFVAPKQNFQQLQDQMLTVCVEIKGALMLPHFNS KAMVIVVNNKGJELVAVRKEQQRQORREEDDEDEEGNSREVRRYRLRKEGD VFIMPAHFVAI.NASSELHLGFGINAENNHRIFLAGDKNDVIDQIEQAKDLAPFGS GSQVEKLI.KNQESHFVSARPOSQSPSSPEKESPEKEDQENQGGKPLLSILKA FN"		
3'UTR	1931..2032		
	/gene="Ara h I"		
polyA_signal	2005..2010		
	/gene="Ara h I"		
polyA_site	2032		
	/gene="Ara h I"		
ORIGIN			
Query Match	100.0%;	Score 2032;	DB 8; Length 2032;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATCAGAGGGAG	60
Db	1	AATAATCATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATCAGAGGGAG	60
QY	61	GGTTTCTCCACTGATGCTGTGCTAGGGATCCCTGTCTGGCTTCAGTTTCTGCAACGCA	120
Db	61	GGTTTCTCCACTGATGCTGTGCTAGGGATCCCTGTCTGGCTTCAGTTTCTGCAACGCA	120
QY	121	TGCCAAGTCATCCTTACAGAGAAACACAGAAACCCCTCGGCCACAGAGTGCTCTCA	180
Db	121	TGCCAAGTCATCCTTACAGAGAAACACAGAAACCCCTCGGCCACAGAGTGCTCTCA	180
QY	181	GAGTTCTCAACAGGACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGCTGCACCAA	240
Db	181	GAGTTCTCAACAGGACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGCTGCACCAA	240

Qy	1321	CCAGCTTCAGGACCTGGACATGATGCTCA	CCTGTGTAGAGATCAAGAGGAGCTTTGAT	1380
Db	1321	CCAGCTTCAGGACCTGGACATGATGCTCA	CCTGTGTAGAGATCAAGAGGAGCTTTGAT	1380
Qy	1381	GCTCCACACATTCACACTCAAGGCCATGG	TTATCGTCGTCAACAAAGGAACTGGAAA	1440
Db	1381	GCTCCACACATTCACACTCAAGGCCATGG	TTATCGTCGTCAACAAAGGAACTGGAAA	1440
Qy	1441	CCTTGAACTCGTGCTGTAAGAAAGAGCA	CAACACAGAGGGGACGGCGGGAAGAGGA	1500
Db	1441	CCTTGAACTCGTGCTGTAAGAAAGAGCA	CAACACAGAGGGGACGGCGGGAAGAGGA	1500
Qy	1501	GGACGAAAGCAAGAGAGGAGGAAAGTAA	CAGAGAGGTCGTAGGTACACAGCGAGTT	1560
Db	1501	GGACGAAAGCAAGAGAGGAGGAAAGTAA	CAGAGAGGTCGTAGGTACACAGCGAGTT	1560
Qy	1561	GAAGGAGGCGATGTGTTTCATCGCAGCA	GCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Db	1561	GAAGGAGGCGATGTGTTTCATCGCAGCA	GCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Qy	1621	CGAACTCCATCTGCTTGGCTTCGGTATC	AAACGCTGAAAAACAACACAGAAATCTTC	1680
Db	1621	CGAACTCCATCTGCTTGGCTTCGGTATC	AAACGCTGAAAAACAACACAGAAATCTTC	1680
Qy	1681	AGGTGATAAGGACAATGTGATAGACCA	GATAGAGCAAGGATTTAGCATTCCTCC	1740
Db	1681	AGGTGATAAGGACAATGTGATAGACCA	GATAGAGCAAGGATTTAGCATTCCTCC	1740
Qy	1741	TGGGTCGGGTGAACAGTTGAGAAGCTC	ATCAAAAACCAAGCAATCTCCTTTGTGAG	1800
Db	1741	TGGGTCGGGTGAACAGTTGAGAAGCTC	ATCAAAAACCAAGCAATCTCCTTTGTGAG	1800
Qy	1801	TGCTCGTCTCAATCTCAATCTCAATCT	CCGTCTCTCGTCTCTCGAAGAGTCTCTG	1860
Db	1801	TGCTCGTCTCAATCTCAATCTCAATCT	CCGTCTCTCGTCTCTCGAAGAGTCTCTG	1860
Qy	1861	AGAGGATCAAGAGGAGGAAAAACCAAG	GAGGAGGTCCTTTTGAAGGC	1920
Db	1861	AGAGGATCAAGAGGAGGAAAAACCAAG	GAGGAGGTCCTTTTGAAGGC	1920
Qy	1921	TTTTAACTGAGATGGAGGCAACTTGTAT	GATCGATATAAGATCACGCTTTGTACT	1980
Db	1921	TTTTAACTGAGATGGAGGCAACTTGTAT	GATCGATATAAGATCACGCTTTGTACT	1980
Qy	1981	CTACTATCCAAAACCTTATCAATAAATA	AAAAACGTTTGTGCGTTGTTCTCC	2032
Db	1981	CTACTATCCAAAACCTTATCAATAAATA	AAAAACGTTTGTGCGTTGTTCTCC	2032
RESULT 4				
AR257470				
LOCUS	AR257470 2041 bp DNA linear PAT 20-DEC-2002			
DEFINITION	Sequence 23 from patent US 6486311.			
ACCESSION	AR257470			
VERSION	AR257470.1 GI:27307483			
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2041)			
AUTHORS	Burke,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A., Helm,R.M. and Bannon,G.A.			
TITLE	Peanut allergens and methods			
JOURNAL	Patent: US 6486311-A 23 26-NOV-2002;			
FEATURES	Location/Qualifiers			
source	1..2041			
	/organism="unknown"			
	/mol_type="genomic DNA"			
ORIGIN				
Query Match	100.0%; Score 2032; DB 6; Length 2041;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 2032; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			

QY	1	AATAATCATATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGGAG	60
Db	1		
QY	61	GGTTTCTCCACATGATGCTGTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA	120
Db	61		
QY	121	TGCCAAGTTCATCACCTTTACAGAGAAACACAGAGAAACCCCTGCGCCACAGAGGTGCTTCCA	180
Db	121		
QY	181	GAGTTGTCTAACAGGAAACCGGATGACTTTGAAGCAAAAGGATGCGAGTCTCGTGTGCAACCA	240
Db	181		
QY	241	GCTCGAGTATGATCCTCGTTGTCTCTATGATCCTCGAGGACACACACACACACCAACCA	300
Db	241		
QY	301	ACGTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTACGATGATGACCG	360
Db	301		
QY	361	CGGTCAACCCGAAAGAGAGGAGGCGATGCGGGAACAGCTGGACCCAGTGGACCCGAGGGAGCGTGA	420
Db	361		
QY	421	AGAGGAGAGAGTGGAGCAACCCAGAGAAAGTGGAGCGGCAAGTTCATCAGCAGCC	480
Db	421		
QY	481	ACGGAATAAGCCCGGAGGAGGAGGAGAAACAAAGAGTGGGGAACACCAAGGTAGCCA	540
Db	481		
QY	541	TGTGAGGAGAGAAACATCTCGGAAACAAACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC	600
Db	541		
QY	601	CGCTACGGGAACCAACCGGTAGGATCGGGTCTGTCAGAGGTTTGACCAAGGTCAAG	660
Db	601		
QY	661	GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCGCAACCTTAACACTCT	720
Db	661		
QY	721	TGTTCTTCCAAAGCAGCTGATGATTAACATCTTGTGTATCCAGCAAGGGCAAGCCAC	780
Db	721		
QY	781	CGTGACCGTAGCAATGCAATAACAGAAAGCTTTAATCTTGACGAGGCGCATGCACT	840
Db	781		
QY	841	CAGAAATCCCATCCGGTTTCATTTCTCATCTTTGAACCGCCATGACAAACAGAACTCAG	900
Db	841		
QY	901	AGTAGCTAAAATCTCCATGCGGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC	960
Db	901		
QY	961	GAGCAGCGGAGACCAATCATCTTGTGAGGCTTCAGCAGGAATACCTTGGAGGCGCC	1020
Db	961		
QY	1021	CTTCAATCGGAATCAATGAGATACGAGGCTGTGTTAGAGAAATGACAGGAGGTGA	1080
Db	1021		

```
QY 1081 GCAAGAGGAGAGAGGCGAGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGAGT 1140
Db 1081 GCAAGAGGAGAGAGGCGAGAGGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGAGT 1140
QY 1141 GATAGTCNAAGTGTCAAAGGAGACGTTGAAGAACTTACTAAGCAGCGCTAAATCGTCTC 1200
Db 1141 GATAGTCNAAGTGTCAAAGGAGACGTTGAAGAACTTACTAAGCAGCGCTAAATCGTCTC 1200
QY 1201 AAAGAAAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTTGAGAGAGGCGA 1260
Db 1201 AAAGAAAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTTGAGAGAGGCGA 1260
QY 1261 GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCACAGCAAGAAAGACC 1320
Db 1261 GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCACAGCAAGAAAGACC 1320
QY 1321 CGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGAT 1380
Db 1321 CGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGAT 1380
QY 1381 GCTCCACACATTTCAACTCAAAAGGCCATGGTTATCGTTCGTCAACAAAGGAACCTGGAAA 1440
Db 1381 GCTCCACACATTTCAACTCAAAAGGCCATGGTTATCGTTCGTCAACAAAGGAACCTGGAAA 1440
QY 1441 CTTTGAACCTCGTGGCTGTGAAGAAAGACAAACAGAGGGGACGGCGGAAGAGAGGA 1500
Db 1441 CTTTGAACCTCGTGGCTGTGAAGAAAGACAAACAGAGGGGACGGCGGAAGAGAGGA 1500
QY 1501 GGNAGAGACGAGAGAGAGGAGGAGTAAACAGAGAGGTTCGTAGGTACACAGCGAGGTT 1560
Db 1501 GGNAGAGACGAGAGAGAGGAGGAGTAAACAGAGAGGTTCGTAGGTACACAGCGAGGTT 1560
QY 1561 GAAGGAAGCGCATGTGTTCAATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
Db 1561 GAAGGAAGCGCATGTGTTCAATCATGCGCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620
QY 1621 CGAATCTCAATCTGTTGGCTTCGGTATCAACGCTGAAACAAACCAAGAGTCTTCCTTGC 1680
Db 1621 CGAATCTCAATCTGTTGGCTTCGGTATCAACGCTGAAACAAACCAAGAGTCTTCCTTGC 1680
QY 1681 AGGTGATAGGACATGTGATAGACAGATAGAGAGCAAGCGAGGATTTAGCATTCCTC 1740
Db 1681 AGGTGATAGGACATGTGATAGACAGATAGAGAGCAAGCGAGGATTTAGCATTCCTC 1740
QY 1741 TGGGTCCGGTGAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCCGGTGAACAAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAG 1800
QY 1801 TGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGTCTGCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTGAGAA 1860
QY 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC 1920
QY 1921 TTTTAACTCAGAAATGAGGCAACTTGTATGTATCGATAAAGATCAAGCTTTTGTACT 1980
Db 1921 TTTTAACTCAGAAATGAGGCAACTTGTATGTATCGATAAAGATCAAGCTTTTGTACT 1980
QY 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032
Db 1981 CTACTATCAAAAACCTTATCAATAAATAAAAAAGTTTGTGCGTTGTTCTCC 2032
```

```
RESULT 5
AX155331
LOCUS AX155331
DEFINITION Sequence 4 from Patent WO0140264.
ACCESSION AX155331
VERSION AX155331.1 GI:14536766
KEYWORDS Arachis hypogaea (peanut)
SOURCE
```

```
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1
REFERENCE
AUTHORS Bannon,G.A., Burks,W.A., Caplan,M.J., Sampson,H. and Sosin,H.
TITLE Peptide antigens
JOURNAL Patent: WO 0140264-A 4 07-JUN-2001;
The University of Arkansas (US)
; MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK
(US)
FEATURES
source
1..2032
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"
ORIGIN
Query Match 99.4%; Score 2019.2; DB 6; Length 2032;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2024; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 AATAATCATATATATTCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
Db 1 AATAATCATATATATTCATCAATCATATATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
QY 61 GGTTCCTCCATCGATGCTGTTGCTAGGGATCCTTGCTCTGCTTCAGTTCTTGCACCGCA 120
Db 61 GGTTCCTCCATCGATGCTGTTGCTAGGGATCCTTGCTCTGCTTCAGTTCTTGCACCGCA 120
QY 121 TGCCAAAGTCATCACCTTACCAAGAGAAAACAGAGAAACCCCTGCGCCAGAGGTGCTCCA 180
Db 121 GGCCAAAGTCATCACCTTACCAAGAGAAAACAGAGAAACCCCTGCGCCAGAGGTGCTCCA 180
QY 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCAACCA 240
Db 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCAACCA 240
QY 241 GCTCGAGTATGATCTCGTTGCTATGATCTCGAGGACACACTGCGACACACCAACCA 300
Db 241 GCTCGAGTATGATCTCGTTGCTATGATCTCGAGGACACACTGCGCGCCACCAACCA 300
QY 301 ACGTTCCCTCTCAGGGAGCGGACACAGTGGCGCCCAACCCGAGAGACTACGATGATGACCG 360
Db 301 ACGTTCCCTCTCAGGGAGCGGACACAGTGGCGCCCAACCCGAGAGACTACGATGATGACCG 360
QY 361 CGGTCAACCCCGAGAGAGGAAGGCGCGATGGGGACAGCTGGACCGAGGAGCGGTGA 420
Db 361 CGGTCAACCCCGAGAGAGGAAGGCGCGATGGGGACAGCTGGACCGAGGAGCGGTGA 420
QY 421 AAGAGAAGAAGACTGGAGACCAACCAAGAGAAAGTTGGAGGCGACCAAGTCTATCAGCAGCC 480
Db 421 AAGAGAAGAAGACTGGAGACCAACCAAGAGAAAGTTGGAGGCGACCAAGTCTATCAGCAGCC 480
QY 481 ACGGAAAAAAGCCCGGAGGAAGAGAAAGAGAAACAAGAGTGGGAAACAACAGGTAGCCA 540
Db 481 ACGGAAAAAAGCCCGGAGGAAGAGAAAGAGAAACAAGAGTGGGAAACAACAGGTAGCCA 540
QY 541 TGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGCGGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGCGGTTTAGCAC 600
QY 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
Db 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
QY 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAAAACCTTAACTCT 720
Db 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGAGATCGAGGCAAAACCTTAACTCT 720
QY 721 TGTTCCTCCAGAGCAGCTGATGCTGATAAATCCTCTGTTATCCAGGAGGCAAGCCAC 780
|||||
```


Db 241 GCTCGAGTATGATCCCTGGTTGTCTATGATCTCGAGGACACACTGGGCACCAACCA 300
QY 301 ACCTTCCCTCCAGGGGAGCGGACAGCTGGCCGCCCAACCCGGAGACTACGATGATGACCG 360
Db 301 ACCTTCCCTCCAGGGGAGCGGACAGCTGGCCGCCCAACCCGGAGACTACGATGATGACCG 360
QY 361 CCCTCAACCCCGAAGAGAGAGAGAGGCGCGATGGGGAACAGCTTGGACCCGAGGGAGCGTGA 420
Db 361 CCCTCAACCCCGAAGAGAGAGAGAGGCGCGATGGGGAACAGCTTGGACCCGAGGGAGCGTGA 420
QY 421 AAGAGAGAGAGACTGGAGCAACCAAGAGAGAGATTGGAGCGGACCAAGTCATCAGCAGCC 480
Db 421 AAGAGAGAGAGACTGGAGCAACCAAGAGAGAGATTGGAGCGGACCAAGTCATCAGCAGCC 480
QY 481 ACGGAAATTAAGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 ACGGAAATTAAGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 TGTGAGGAGAGAAACATCTCGGAAACAACTTCTTACTTCCCGTCAAGGCGGTTAGCAC 600
Db 541 TGTGAGGAGAGAAACATCTCGGAAACAACTTCTTACTTCCCGTCAAGGCGGTTAGCAC 600
QY 601 CCGCTACGGGACCAAAACCGTAGGATCCGGTCTCGAGAGGTTTGACCAAGGTCAG 660
Db 601 CCGCTACGGGACCAAAACCGTAGGATCCGGTCTCGAGAGGTTTGACCAAGGTCAG 660
QY 661 GCAGTTTCAGAACTCTCAGAACTCACCGTATTGTGAGATCGAGGCCAAACCTTAACTCT 720
Db 661 GCAGTTTCAGAACTCTCAGAACTCACCGTATTGTGAGATCGAGGCCAAACCTTAACTCT 720
QY 721 TGTCTTCCCAAGCAGCGTGATGCTGATTAACATCTTGTATTCCAGCAAGGCGCAAGCCAC 780
Db 721 TGTCTTCCCAAGCAGCGTGATGCTGATTAACATCTTGTATTCCAGCAAGGCGCAAGCCAC 780
QY 781 CGTGACGCTAGCAAAATGCAATTAACAGAGAGCTTTAATCTTGACGAGGCGCATGCACT 840
Db 781 CGTGACGCTAGCAAAATGCAATTAACAGAGAGCTTTAATCTTGACGAGGCGCATGCACT 840
QY 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAACGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAACGAACTCAG 900
QY 901 AGTAGTAAATCTCCATCCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC 960
Db 901 AGTAGTAAATCTCCATCCCGTTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGGC 960
QY 961 GAGCAGCCGAGCAACATCTACTTGTGAGGCTTTCAGCAGGATACGTTGGAGGCGGC 1020
Db 961 GAGCAGCCGAGCAACATCTACTTGTGAGGCTTTCAGCAGGATACGTTGGAGGCGGC 1020
QY 1021 CTTCAATCGGGAATTCATAGATACGAGGCTGCTGTGTAGAGAGATGCGAGGAGTGA 1080
Db 1021 CTTCAATCGGGAATTCATAGATACGAGGCTGCTGTGTAGAGAGATGCGAGGAGTGA 1080
QY 1081 GCAAGAGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT 1140
Db 1081 GCAAGAGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT 1140
QY 1141 GATAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTTAGCAGCGCTAAATCGTCTC 1200
Db 1141 GATAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTTAGCAGCGCTAAATCGTCTC 1200
QY 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATACCAACCCCAATCACTTGAGAGAGGCGA 1260
Db 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATACCAACCCCAATCACTTGAGAGAGGCGA 1260
QY 1261 GCCCGATCTTTCTAAACAATTTGGGAGGTTATTGAGGTGAAGCCAGACAAGAAACCC 1320
Db 1261 GCCCGATCTTTCTAAACAATTTGGGAGGTTATTGAGGTGAAGCCAGACAAGAAACCC 1320
QY 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGAT 1380

QY 1381 GCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTCGTCAACAAAGGAACCTGGAAA 1440
Db 1381 GCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTCGTCAACAAAGGAACCTGGAAA 1440
QY 1441 CTTTGAATCTCGTGGCTGTAAAGAAAGAGACAACAGAGGGGACCGCGGGAAGAGAGGA 1500
Db 1441 CTTTGAATCTCGTGGCTGTAAAGAAAGAGACAACAGAGGGGACCGCGGGAAGAGAGGA 1500
QY 1501 GGACCAAGACGGAAGAGAGGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
Db 1501 GGACCAAGACGGAAGAGAGGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
QY 1561 GAAAGGAGCGCATGTGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCTC 1620
Db 1561 GAAAGGAGCGCATGTGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCTC 1620
QY 1621 CGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACAGAAATCTTCTCTTGC 1680
Db 1621 CGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACAGAAATCTTCTCTTGC 1680
QY 1681 AGGTGATAGGACAAATGTGTAGACCAATAGAGAGCAAGCGAAGGATTTAGCATTCCTC 1740
Db 1681 AGGTGATAGGACAAATGTGTAGACCAATAGAGAGCAAGCGAAGGATTTAGCATTCCTC 1740
QY 1741 TGGGTCGGGTGAACAAAGTTGAGAACTCATCAAAACCCAGAAAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCGGGTGAACAAAGTTGAGAACTCATCAAAACCCAGAAAGGAATCTCACTTTGTGAG 1800
QY 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTCGAGAAAGAGTCTCTCGAGAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTCGAGAAAGAGTCTCTCGAGAA 1860
QY 1861 AGAGCATCAAGAGGAGGAGAAACCAAGGAGGAAACCAAGGAGGAGGTCCTTCAATTTTCAAGGC 1920
Db 1861 AGAGCATCAAGAGGAGGAGAAACCAAGGAGGAGGAAACCAAGGAGGAGGTCCTTCAATTTTCAAGGC 1920
QY 1921 TTTTAACTGA 1930
Db 1921 TTTTAACTGA 1930

RESULT 7
AR257463
LOCUS AR257463 1930 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6486311.
ACCESSION AR257463
VERSION AR257463.1 GI:27307476
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1930)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannon,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 3 26-NOV-2002;
FEATURES Location/Qualifiers
source 1..1930
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 95.0%; Score 1930; DB 6; Length 1930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
Db 1 AATATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG 60
QY 61 GGTTCCTCAGTATGCTGTTGCTAGGATCCTTGTCTCGCTCAGTTCTTGCACCGA 120
Db 61 GGTTCCTCAGTATGCTGTTGCTAGGATCCTTGTCTCGCTCAGTTCTTGCACCGA 120

Db 61 GGTTCCTCCATGATGCTGTGCTAGGATCCTTGTCTCTGGCTTCAGTTTCTGCAAGCA 120
Qy 121 TGCAGATGATACCTTTACAGAGAAACAGAGAAACCCCTGCGCCACAGAGTGCTCCA 180
Db 121 TGCAGATGATACCTTTACAGAGAAACAGAGAAACCCCTGCGCCACAGAGTGCTCCA 180
Qy 181 GAGTGTCAACAGAAACCGGATGCTTGAAGCAAAAGGATGCGATCTCGCTGCAACCA 240
Db 181 GAGTGTCAACAGAAACCGGATGCTTGAAGCAAAAGGATGCGATCTCGCTGCAACCA 240
Qy 241 GCTCGAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTGGCAACCAACCA 300
Db 241 GCTCGAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTGGCAACCAACCA 300
Qy 301 ACCTTCCCTCCAGGGAGCGGACAGTGGCCGCCAACCCGGAGACTACGATGATGACCG 360
Db 301 ACCTTCCCTCCAGGGAGCGGACAGTGGCCGCCAACCCGGAGACTACGATGATGACCG 360
Qy 361 CCCTCAACCCGAGAGAGAGGAGGCGGATGGGACCGAGCTGGACCGAGGGAGCGTGA 420
Db 361 CCCTCAACCCGAGAGAGAGGAGGCGGATGGGACCGAGCTGGACCGAGGGAGCGTGA 420
Qy 421 AAGAGAGAGAGACTGGAGACCAACCAAGAGAGAGATGGAGCGGACCAAGTCAACAGC 480
Db 421 AAGAGAGAGAGACTGGAGACCAACCAAGAGAGAGATGGAGCGGACCAAGTCAACAGC 480
Qy 481 ACAGAAATAGGCCCGAG 540
Db 481 ACAGAAATAGGCCCGAG 540
Qy 541 TGTGAGGAGAGAACTCTCGGAACACCTTTCTACTTCCGTCAAGGGGTTAGCAC 600
Db 541 TGTGAGGAGAGAACTCTCGGAACACCTTTCTACTTCCGTCAAGGGGTTAGCAC 600
Qy 601 CCCTTACGGGAAACAAAACGGTAGGATCCGGGTCCTGAGAGGTTTGAACAAAGGTCAAG 660
Db 601 CCCTTACGGGAAACAAAACGGTAGGATCCGGGTCCTGAGAGGTTTGAACAAAGGTCAAG 660
Qy 661 GCAGTTTCAGATCTCCAGAAACACCGTATTTGACAGATCGAGGCCAAACCTTAACTCT 720
Db 661 GCAGTTTCAGATCTCCAGAAACACCGTATTTGACAGATCGAGGCCAAACCTTAACTCT 720
Qy 721 TGTCTTCCAGACGCTGATGCTGATACATCTTGTATCCAGCAAGGGCAAGCCAC 780
Db 721 TGTCTTCCAGACGCTGATGCTGATACATCTTGTATCCAGCAAGGGCAAGCCAC 780
Qy 781 CGTGACCGTAGCAATCGGCAATTAACAGAAAGAGCTTTAATCTTGACAGGGCCATGCACT 840
Db 781 CGTGACCGTAGCAATCGGCAATTAACAGAAAGAGCTTTAATCTTGACAGGGCCATGCACT 840
Qy 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGGCCATGACCAACAGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGGCCATGACCAACAGAACTCAG 900
Qy 901 AGTAGTAAATCTCCATGCCGTTTAAACACCGGCCAGTTTGAGATTTCTTCCGGC 960
Db 901 AGTAGTAAATCTCCATGCCGTTTAAACACCGGCCAGTTTGAGATTTCTTCCGGC 960
Qy 961 GAGCAGCGAGACCAATCATCTTCTTACAGGGGTTTACAGAGAAATACGTTGGAGGCCG 1020
Db 961 GAGCAGCGAGACCAATCATCTTCTTACAGGGGTTTACAGAGAAATACGTTGGAGGCCG 1020
Qy 1021 CTTCAATCGGAAATCAATGAGATACGAGGGTGTCTTTAGAGAGAAATGCAAGAGGTGA 1080
Db 1021 CTTCAATCGGAAATCAATGAGATACGAGGGTGTCTTTAGAGAGAAATGCAAGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGCGAGAGGCGATGGAGTACTCGAGTACTGAGAAACAAAGAGGACT 1140
Db 1081 GCAAGAGGAGAGGCGAGAGGCGATGGAGTACTCGAGTACTGAGAGAAACAAAGAGGAGT 1140
Qy 1141 GATAGTCAAAAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 GATAGTCAAAAGTCAAAAG 1200

RESULT 8

BD107898

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BD107898
Methods and reagents for decreasing allergic reactions.

BD107898
JP 2002501748-A/1.

unidentified

unclassified.

1 (bases 1 to 1930)

Sosin, H., Banon, G.A., Jr, W.A.B. and Samphson, H.A.

Methods and reagents for decreasing allergic reactions

Patent: JP 2002501748-A 1 22-JAN-2002;

UNIVERSITY OF ARKANSAS, MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY

UNIVERSITY OF NEW YORK, HOWARD SOSIN

OS Arachis L. (Peanut)

PN JP 2002501748-A/1

PD 22-JAN-2002

1930 bp DNA linear PAT 18-SBP-2002

```
PF 29-JAN-1999 JP 2000529437
PR 31-JAN-1998 US 60/073283,13-FEB-1998 US 60/074590 PR
13-FEB-1998 US 60/074624,13-FEB-1998 US 60/074633 PR
27-AUG-1998 US 09/141220
PI HAWADO SOSIN, GARY A BANON, WESLEY A BIRX JR, HYU A SAMPTHON PC
C12N15/09, A01H5/00, A01K67/027, A61K39/35, C07K14/37, C07K14/415, PC
C07K14/435,
PC C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/00, C12N15/00, C12N5/
PC 00
CC Methods and reagents for decreasing allergic reactions PH
Key Location/Qualifiers
FT source 1..1930
FT /organism='Arachis l. (Peanut)'.
FEATURES
source Location/Qualifiers
1..1930
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 95.0%; Score 1930; DB 6; Length 1930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGGAG 60
DB 1 AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGGAG 60
QY 61 GGTTCCTCACATGATGCTGTGCTAGGATCCTTGCTGGCTTCAGTTTCTGCAACGCA 120
DB 61 GGTTCCTCACATGATGCTGTGCTAGGATCCTTGCTGGCTTCAGTTTCTGCAACGCA 120
QY 121 TCCCAAGTCATCACCTTACAGAAGAAACAGAGAACCCCTCGCCCGAGAGTGCTTCCA 180
DB 121 TCCCAAGTCATCACCTTACAGAAGAAACAGAGAACCCCTCGCCCGAGAGTGCTTCCA 180
QY 181 GAGTTGTCAACAGGAAACCGGATGACTTGAGCAAAAGGCATCGAGTCTCGCTGCAACAA 240
DB 181 GAGTTGTCAACAGGAAACCGGATGACTTGAGCAAAAGGCATCGAGTCTCGCTGCAACAA 240
QY 241 GCTCAGATGATGATCCTCGTTGTCTATGATCTCTGAGGACACATGCGACCAACCA 300
DB 241 GCTCAGATGATGATCCTCGTTGTCTATGATCTCTGAGGACACATGCGACCAACCA 300
QY 301 AGTTTCCTCCAGGGAGCGACACGTGGCGCCAAACCCGAGACTAGATGATGACCG 360
DB 301 AGTTTCCTCCAGGGAGCGACACGTGGCGCCAAACCCGAGACTAGATGATGACCG 360
QY 361 CCGTCAACCCGAAAGAGAGGAAAGGCGCGATGGGGACCGAGTGGACCGAGGGACGTGA 420
DB 361 CCGTCAACCCGAAAGAGAGGAAAGGCGCGATGGGGACCGAGTGGACCGAGGGACGTGA 420
QY 421 AAGAGAGNAGACTGGAGACACCAAGAGAGATTTGGAGGGACCAAGTCTACAGCAGCC 480
DB 421 AAGAGAGNAGACTGGAGACACCAAGAGAGATTTGGAGGGACCAAGTCTACAGCAGCC 480
QY 481 ACGGAAATTAAGCCCGGAGGAGAGAGAGGAAACAAAGATGGGAAACACAGGTAGCCA 540
DB 481 ACGGAAATTAAGCCCGGAGGAGAGAGAGGAGAAACAAAGATGGGAAACACAGGTAGCCA 540
QY 541 TGTGAGGAGAAACATCTCGGAAACAAACCTTTCTACTTTCCGTCAAGCGGTTTAGCAC 600
DB 541 TGTGAGGAGAAACATCTCGGAAACAAACCTTTCTACTTTCCGTCAAGCGGTTTAGCAC 600
QY 601 CCGCTACGGGAAACCAACGGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAG 660
DB 601 CCGCTACGGGAAACCAACGGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAG 660
QY 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCAAAACCTTAACTCT 720
DB 661 GCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCAAAACCTTAACTCT 720
QY 721 TGTTCCTTCCCAAGCAGGTGATGCTGATAACATCCTTGTATTCCAGCAAGGCAAGCCAC 780
```

```
DB 721 TGTTCCTTCCCAAGCAGGTGATGCTGATAACATCCTTGTATTCCAGCAAGGCAAGCCAC 780
QY 781 CTTGACCCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 840
DB 781 CTTGACCCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 840
QY 841 CAGAAATCCATCCGTTTCATTTCTACATCTTTGAACCGCCATGACAAACCAAGAACTCAG 900
DB 841 CAGAAATCCATCCGTTTCATTTCTACATCTTTGAACCGCCATGACAAACCAAGAACTCAG 900
QY 901 AGTAGCTAAATCTCATGCCCCGTTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCC 960
DB 901 AGTAGCTAAATCTCATGCCCCGTTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCC 960
QY 961 GAGCAGCCGAGACCAATCATCTTCTAGGCTTTCAGAGGATTAAGTTGGAGGCGC 1020
DB 961 GAGCAGCCGAGACCAATCATCTTCTAGGCTTTCAGAGGATTAAGTTGGAGGCGC 1020
QY 1021 CTTCAATGCGGAAATTCATGAGATACGAGAGGTGCTTTAGAGAGAAATGCAAGAGGTGA 1080
DB 1021 CTTCAATGCGGAAATTCATGAGATACGAGAGGTGCTTTAGAGAGAAATGCAAGAGGTGA 1080
QY 1081 GCAAGAGAGAGAGGCGCAGCGGATGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT 1140
DB 1081 GCAAGAGAGAGAGGCGCAGCGGATGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGT 1140
QY 1141 GATAGTCAAGGTGCAAGGAGGACGTTGAAGAACTTACTAAGCAACGCTTAATCCGCTCTC 1200
DB 1141 GATAGTCAAGGTGCAAGGAGGACGTTGAAGAACTTACTAAGCAACGCTTAATCCGCTCTC 1200
QY 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACCACCAATCAACTTTGAGAGAGGCGA 1260
DB 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACCACCAATCAACTTTGAGAGAGGCGA 1260
QY 1261 GCCGATCTTTCTAACTTTGGGAGTATTTGAGGTGAAGCCAGACAAGAAACCC 1320
DB 1261 GCCGATCTTTCTAACTTTGGGAGTATTTGAGGTGAAGCCAGACAAGAAACCC 1320
QY 1321 CCAGTTCCAGACCTGGACATGCTCCTCGTGTAGAGATCAAGAGAGGCTTTGAT 1380
DB 1321 CCAGTTCCAGACCTGGACATGCTCCTCGTGTAGAGATCAAGAGAGGCTTTGAT 1380
QY 1381 GCTCCCACTTTCAACTCAAAAGGCGCATGTTTATCGTCTCGTCAACAAAGAACTGGAAA 1440
DB 1381 GCTCCCACTTTCAACTCAAAAGGCGCATGTTTATCGTCTCGTCAACAAAGAACTGGAAA 1440
QY 1441 CTTTGAACTCTGTGCTGTAAGAAAGAGCAACACAGAGGGGACGCGGGAGAGAGGA 1500
DB 1441 CTTTGAACTCTGTGCTGTAAGAAAGAGCAACACAGAGGGGACGCGGGAGAGAGGA 1500
QY 1501 GGACGAGACCAAGAGAGGAGGAGTAACAGAGAGGTGAGTACAGAGGAGTTCAGAGGAGTT 1560
DB 1501 GGACGAGACCAAGAGAGGAGGAGTAACAGAGAGGTGAGTACAGAGGAGTTCAGAGGAGTT 1560
QY 1561 GAAGAGAGGCGATGTTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
DB 1561 GAAGAGAGGCGATGTTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
QY 1621 CGAACTCATCTGCTTGGCTTCCGATTCAGCGTGAAACCAACCAAGAAATCTTCTTTCG 1680
DB 1621 CGAACTCATCTGCTTGGCTTCCGATTCAGCGTGAAACCAACCAAGAAATCTTCTTTCG 1680
QY 1681 AGGTGATAAGGACCAATGTGATAGACCAAGAGAGGAGTTCAGAGGAGTTCAGAGGAGTT 1740
DB 1681 AGGTGATAAGGACCAATGTGATAGACCAAGAGAGGAGTTCAGAGGAGTTCAGAGGAGTT 1740
QY 1741 TGGGTTCGGGTGAAACAAAGTTGAGAGGCTCATCAAAAAACAGAGGAAATCTCACTTTGTGAG 1800
DB 1741 TGGGTTCGGGTGAAACAAAGTTGAGAGGCTCATCAAAAAACAGAGGAAATCTCACTTTGTGAG 1800
QY 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCT 1860
```



```
QY 1561 GAAGGAAGCGATGTGTTTCATCATGCGAGCAGCTCATCCAGTAGGCATCAACGCTTCCTC 1620
Db 1535 GAAGGAAGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATCAACGCTTCCTC 1594
QY 1621 CGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAGAAACAAACAGAAATCTTCCTTGC 1680
Db 1595 CGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAGAAACAAACAGAAATCTTCCTTGC 1654
QY 1681 AGTGATAGGACAAATGTGATAGACAGATAGAGAGCAAGCAAGAGATTTAGCAATTCCTC 1740
Db 1655 AGTGATAGGACAAATGTGATAGACAGATAGAGAGCAAGCAAGAGATTTAGCAATTCCTC 1714
QY 1741 TGGGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGAAATCTCACTTTGTGAG 1800
Db 1715 TGGTTCGGGTGAACAAAGTTGAGAAGCTCATCAAAACCAAGAGAGTCTCACTTTGTGAG 1774
QY 1801 TGCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAAGAGATCTCTCGAGAA 1860
Db 1775 TGCTGCTCTCAATCTC-----CGTCTCTCTGAAAA 1806
QY 1861 AGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC 1920
Db 1807 AGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCACCTCTTCAATTTTGAAGGC 1866
QY 1921 TTTAACTGAGAAATGGAGCAACTTGTATGATATCGATAAAGATCAAGCTTTTGTACT 1980
Db 1867 TTTAACTGAGAAATGGAGAACTTGTATGATATCATTAATTAAGATCAAGCTTTTGTAA 1926
QY 1981 CTACTATCCAAAACCTTATCAATAATAAAGCTTTGTGCTGCTTCTCC 2032
Db 1927 CTACTATCCAAAACCTTATCAATAATAAAGCTTTGTGCTGCTTCTCC 1978

RESULT 10
AR257467
LOCUS AR257467 1949 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 19 from patent US 6486311.
ACCESSION AR257467
VERSION AR257467.1 GI:27307480
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1949)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannon,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 19 26-NOV-2002;
FEATURES
source
1..1949
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 86.2%; Score 1752.4; DB 6; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;

QY 48 CAATGAGAGGAGGGTTTCTCCACTGATGCTGCTGCTAGGATCCTTGTCTGGCTTCAG 107
Db 1 CAAATGAGAGGAGGGTTTCTCCACTGATGCTGCTGCTGGATCCTTGTCTGGCTTCAG 60
QY 108 TTTCTGCAACGCATCCCAAGTATCATCCATTACAGAAAGAACAGAGAACCCCTCGGCC 167
Db 61 TTTCTGCAACGCAGGCCCAAGTACC-----TTACCGGAAACAGAGAACCCCTCGGCC 114
QY 168 AGAGTGTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCCGAGT 227
Db 115 AGAGTGTCTCCAGAGTTGTCAACAGGAACCGGACGACTTGAAGCAAAAGGCATCCGAGT 174
QY 228 CTCGTGCAACCAAGCTCAGATATGATCTCTGTTGTCTATGATCCTCGAGGACACACTG 287
Db 175 CTCGTGCAACCAAGCTCAGATATGATCTCTGTTGTCTATG-----ACACTG 222
```

```
QY 288 GCACCACCAACCAACGTTCCCTTCAGGGAGCGSACACGTGGCCGCCAACCCCGAGACT 347
Db 223 GCGCCACCAACCAACGTCACCTCCAGGGAGCGSACACGTGGCCGCCAACCCCGAGACT 282
QY 348 ACATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGATGGGGACACAGCTGGAC 407
Db 283 ACATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGATGGGGACACAGCTGGAC 342
QY 408 CGAGGAGCGTGAAGAGAGAAAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCA 467
Db 343 CGAGGAGCGTGAAGAGAGAAAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCA 402
QY 468 GTCATCAGCAGCCACGGAATAAGGCCCGAAGAGAGAGAGAGAAACCAAGTGGGAA 527
Db 403 GTCATCAGCAGCCACGGAATAAGGCCCGAAGAGAGAGAGAGAAACCAAGTGGGAA 462
QY 528 CACAGGTAGCATGTGAGGAGAAACATCTCGGAAACAAACCTTTTCTACTTCCCGTCAA 587
Db 463 CACAGGTAGCATGTGAGGAGAAACATCTCGGAAACAAACCTTTTCTACTTCCCGTCAA 522
QY 588 GCGGTTTTAGCACCGCTACCGGAAACCAAAACGCTAGGATCGGGTCTCGCAGAGTTTG 647
Db 523 GCGGTTTTAGCACCGCTACCGGAAACCAAAACGCTAGGATCGGGTCTCGCAGAGTTTG 582
QY 648 ACCAAGGTCAAGGCAGTTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCA 707
Db 593 ACCAAGGTCAAGGCAGTTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCA 642
QY 708 AACCTAACACTCTTTGTTCTTCCCAAGCAGCTGATGTGTGATAACATCTTGTATCCAGC 767
Db 643 GACCTAACACTCTTTGTTCTTCCCAAGCAGCTGATGTGTGATAACATCTTGTATCCAGC 702
QY 768 AAGGCAAGCCACCGTAGCCGTAGCAATGCAATTAACAGAAAGAGCTTTTATCTTGAGC 827
Db 703 AAGGCAAGCCACCGTAGCCGTAGCAATGCAATTAACAGAAAGAGCTTTTATCTTGAGC 762
QY 828 AGGCGCATGCACTCAGAAATCCATCCGTTTTCATTCTCTACATCTTGAACCCCATGACA 887
Db 763 AGGCGCATGCACTCAGAAATCCATCCGTTTTCATTCTCTACATCTTGAATCGACATGACA 822
QY 888 ACCAGAACTCAGAGTAGCTAAATCTCCATCCCGTTTAAACACACCCCGCAGTTTGAGG 947
Db 823 ACCAGAACTCAGAGTAGCTAAATCTCCATGCGCTTAAACACACCCCGCAGTTTGAGG 882
QY 948 ATTTCTTCCCGCGCAGCGCGAGCCAGACCAATCATCTTCTTGCAGGGCTTCAGCAGGATA 1007
Db 883 ATTTCTTCCCGCGCAGCGCGAGCCAGACCAATCATCTTCTTGCAGGGATTCAGCAGGATA 942
QY 1008 CGTTGGAGGCGCCTTCAATCGGAAATTCATATGAGATACGAGGGTGTCTTTAGAGAGA 1067
Db 943 CTTTGGAGGCGCCTTCAATCGGAAATTCATATGAGATACGAGGGTGTCTTTAGAGAGA 1002
QY 1068 ATGAGGAGGTGAGCAAGAGAGAGAGGCGAGCGGATGAGTACTCGGAGTAGTGAGA 1127
Db 1003 ATGAGGAGGTGAGCAAGAGAGAGAGGCGAGCGGATGAGTACTCGGAGTAGTGAGA 1059
QY 1128 ACAATGAGGAGTATAGTCAAGTGTCAAGAGGACGTTTGAAGAACTTACTAAGCAGC 1187
Db 1060 ATAATGAGGAGTATAGTCAAGTGTCAAGAGGACGTTTCAAGAACTTACTAAGCAGC 1119
QY 1188 CTAATCCGTCTCAAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCAATCAACT 1247
Db 1120 CTAATCCGTCTCAAGAAAGGCTCC---GAAAGAGAGATATCACCAACCAATCAACT 1176
QY 1248 TGAGAGAGGCGAGCCCGATCTTTCTAAACATTTTGGGAAGTTTATGAGGTGAAGCCAG 1307
Db 1177 TGAGAGATGGGAGCCCGATCTTTCTAAACATTTTGGGAGTTTATTTGAGGTGAAGCCAG 1236
QY 1308 ACAGAGAAACCCCGAGTTTCAGGACCTGGACATGATGCTCACCTGTGTAGATCAAG 1367
Db 1237 ACAGAGAAACCCCGAGTTTCAGGACCTGGACATGATGCTCACCTGTGTAGATCAAG 1296
```

```
QY 1368 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACA 1427
Db 1297 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACA 1356
QY 1428 AAGGAACTGGAACCTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGC 1487
Db 1357 AAGGAACTGGAACCTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGC 1416
QY 1488 GGGAA-----GAAGAGGAGGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1538
Db 1417 GGGAAACAAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1476
QY 1539 TGGGTAGGTACACAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1598
Db 1477 TGGGTAGGTACACAGGAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1536
QY 1599 CAGTAGGCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGGCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1596
QY 1659 ACAACACAGAAATCTTCTTGAAGGTGATGAAGCAATGTGATAGACCATGAGAGGAGG 1718
Db 1597 ACAACACAGAAATCTTCTTGAAGGTGATGAAGCAATGTGATAGACCATGAGAGGAGG 1656
QY 1719 AAGCGAAGGATTTAGCATTCCTCGGCTCGGCTGCAAGTTGAGAGGAGGAGGAGGAGGAGG 1778
Db 1657 AAGCGAAGGATTTAGCATTCCTCGGCTCGGCTGCAAGTTGAGAGGAGGAGGAGGAGGAGG 1716
QY 1779 AGAAGGAATCTCACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCCGCTCTC 1838
Db 1717 AGAGGAGTCTCACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCCGCTC----- 1765
QY 1839 CTGAGAAAGATCTCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1898
Db 1766 -----GTCTCTGAAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1815
QY 1899 CACTCTCTCAATTTGAGGCTTTAACTGAGAAATGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1958
Db 1816 CACTCTCTCAATTTGAGGCTTTAACTGAGAAATGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1875
QY 1959 AATAAGATCAAGCTTTGTACTCTACTATCAATAAACTATCAATAAACTATCAATAAACTTTG 2018
Db 1876 AATAAGATCAAGCTTTGTAACTACTATCAATAAACTATCAATAAACTATCAATAAACTTTG 1935
QY 2019 TGGCTTTGTTCTCC 2032
Db 1936 TGGCTTTGTTCTCC 1949

RESULT 11
ARQARAH
LOCUS
DEFINITION
Arachis hypogaea (clone P17) Ara h I mRNA, linear, PLN 10-JAN-1995
ACCESSION
L38853
VERSION
L38853.1 GI:620024
KEYWORDS
peanut hypersensitivity.
SOURCE
Arachis hypogaea
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 1949)
Burke,A.W., Cockrell,G., Stanley,J.S., Helm,R.M. and Bannion,G.A.
Recombinant peanut allergen Ara h I expression and IgE binding in
patients with peanut hypersensitivity
Unpublished (1994)
JOURNAL
COMMENT
Original source text: Arachis hypogaea (strain Florunner) Seed cDNA
to mRNA.
FEATURES
Location/Qualifiers
1..1949
/organism="Arachis hypogaea"
/mol_type="mRNA"
```

```
/strain="Florunner"
/db_xref="taxon:3818"
/dev_stage="Seed"
1..1949
/gene="Ara h I"
1..2
/gene="Ara h I"
3..1847
/gene="Ara h I"
/codon_start=1
/protein_id="AAA60336.1"
/db_xref="GI:620025"
/translation="MRGRVSLPLMLLLGLVLSVSAATQAKSPYRKTEPNCAPCLQSC
QEPDLDLKAKESRCTKLEYDPRCYDGTGNQRPGRTRGROPQGYDDDDRRQPR
REGGRWGPAEPREEREDWRQPRDWRPQRPQRKIPRPGREGQEGWGTGTSVPR
ETSPNNPFYFPRFSTRYNGNRIKRVLFQDRSKQFQNLQNHRIYQIEARPNL
VLPKADADNLIYVQGGATVTVANGNNKSFNLDEGHALRIPSGTISVILNRHDNQ
LRVAKISMEVNTPGQFDFPAPSSRDQSYLOGFSRNTLEAAPAFNFIIRRVLLSEN
AGGEQERGORBERTSSDNEGVIKVSKHEVQELTKHAKSVSKGSEEDLTNPIL
RGEEDLNNFGLFEVPDKNPQLQDLMLTLCVEIKEGALMLPFSKAMVIVV
NKGTLNLELVAVRKKQQRREQWEEDEEBEGSNREVRRTARLKGGDVFIIMP
AAHPVAINASSSELHLLGFINAENNRHIFLAGDKDNVIDQIEKQADLAFPGSGEQVE
KLIKQNRSHFVSARPOSQSPSPSEKEDQEEENQCGKPLLSLKA FN"
1848..1949
3' UTR
/gene="Ara h I"
1949
polyA_site
/gene="Ara h I"
ORIGIN
Query Match 86.2%; Score 1752.4; DB 8; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;
QY 48 CAATGAGAGGAGGGTTTCTCCACTGATGCTGTGTCTAGGGATCTCTGTCTGGCTTCAG 107
Db 1 CAATGAGAGGAGGGTTTCTCCACTGATGCTGTGTCTGGATCTCTGTCTGGCTTCAG 60
QY 108 TTTCTGCAACGATGCGCAAGTCACTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167
Db 61 TTTCTGCAACGAGGCGCAAGTCACC-----TTACCGGAGAGAGAGAGAGAGAGAGAGAG 114
QY 168 AGAGGTGCTCCAGAGTGTGTCAACAGGAGACCGAGTGTGTGAAGCAAGAGGAGGAGT 227
Db 115 AGAGGTGCTCCAGAGTGTGTCAACAGGAGACCGAGTGTGAAGCAAGAGGAGGAGT 174
QY 228 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTCTATGATCTCTCGAGGACACACTG 287
Db 175 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTCTATG-----ACACTG 222
QY 288 GCACCAACCAACGATTTCCCTCCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 347
Db 223 GCGCCACCAACCAAGTTCACCTCCAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 282
QY 348 AGCATGATGACCGCGCTCAACCCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 407
Db 283 AGCATGATGACCGCGCTCAACCCCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 342
QY 408 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 467
Db 343 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402
QY 468 GTTCATCAGCAGCCACGGAAAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 527
Db 403 GTTCATCAGCAGCCACGGAAAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGTGG 462
QY 528 CACCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 587
Db 463 CACCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 522
QY 588 GCGCGTTTATGACACCGCTACCGGAGACCAAAACGGTAGGATCGGGGTCTCTCGAGAGGTTG 647
Db 523 GCGCGTTTATGACACCGCTACCGGAGACCAAAACGGTAGGATCGGGGTCTCTCGAGAGGTTG 582
```

QY 648 ACCAAAGGTCAAGGCGATTTTCAGAAATCTCCAGAAATCACCGTATTGTGCGAGTCGAGGCCA 707
Db 583 ACCAAAGGTCAAGGCGATTTTCAGAAATCTCCAGAAATCACCGTATTGTGCGAGTCGAGGCCA 642
QY 708 AACCTAACACTCTGTTCTTCCCAAGCAGCGCTGATGATTAACATCTCTGTTATCCAGC 767
Db 643 GACCTAACACTCTGTTCTTCCCAAGCAGCGCTGATGATTAACATCTCTGTTATCCAGC 702
QY 768 AAGGGCAAGCCACCGTGAACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGAGC 827
Db 703 AAGGACAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGAGC 762
QY 828 AGGGCATGCACTCAGAAATCCATCGGTTTCATTTCTCACTCTTGAAACCCCATGACA 887
Db 763 AGGGCATGCACTCAGAAATCCATCGGTTTCATTTCTCACTCTTGAAATGACATGACA 822
QY 888 ACCAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCGGCCAGTTTGAGG 947
Db 823 ACCAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCGGCCAGTTTGAGG 882
QY 948 ATTTCTCCCGGCGAGCAGCCAGACCAATCATCTTGTGAGGCTTTCAGCAGGAATA 1007
Db 883 ATTTCTCCCGGCGAGCAGCCAGACCAATCATCTTGTGAGGATTCAGCAGGAATA 942
QY 1008 CTTTGAGGCCCGCTTCAATCGGAAATCAATGAGATACGGAGGTGCTGTTAGAGAGA 1067
Db 943 CTTTGAGGCCCGCTTCAATCGGAAATCAATGAGATACGGAGGTGCTGTTAGAGAGA 1002
QY 1068 ATGCAAGGAGTACGAAAGGAGAGGCGCAGCGGATGGAGTACTCGGAGTAGTGAGA 1127
Db 1003 ATGCAAGGAGTACGAAAGGAGAGGCGCAGCGGATGGAGTACTCGGAGTAGTG--- 1059
QY 1128 ACAATGAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCAGCTTCAAGAACTTACTAAGCAGC 1187
Db 1060 ATAATGAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCAGCTTCAAGAACTTACTAAGCAGC 1119
QY 1188 CTAATTCGTTCAAGAAAGGCTCCGAAGAGAGGAGATATCAACCAACCAATCAACT 1247
Db 1120 CTAATTCGTTCAAGAAAGGCTCC--GAAGAGGAAGATATCAACCAACCAATCAACT 1176
QY 1248 TCAGAGAGGCGAGCCCGATCTTTCTAACACTTTGGGAGGTATTGAGGTGAAGCCAG 1307
Db 1177 TGAGAGATGGGAGCCCGATCTTTCTAACACTTTGGGAGGTATTGAGGTGAAGCCAG 1236
QY 1308 ACAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG 1367
Db 1237 ACAAGAGNACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG 1296
QY 1368 AAGGAGCTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCGTCACA 1427
Db 1297 AAGGAGCTTGATGCTCCACACTTCAACTCAAGGCCATGGTCATCGTCGTCGTCACA 1356
QY 1428 AAGGAGCTGGAACTTGAATCTCGTGGCTGTGAAGAAAGAGCAACACAGAGGGACCGC 1487
Db 1357 AAGGAGCTGGAACTTGAATCTCGTGGCTGTGAAGAAAGAGCAACACAGAGGGACCGC 1416
QY 1488 GGGAA-----GAAGAGGAGCAGCAAGCAAGAGAGAGGAGGAGAGTAAACAGAGG 1538
Db 1417 GGGAAACAGATGGGGAAGAGAGGAGGAGAGATGAAGAGAGAGGAGGAGTAAACAGAGG 1476
QY 1539 TCGGTAGGTACACAGGAGTTGAAGGAGGCGATGTGTTTCATATGCGCAGCGTCATC 1598
Db 1477 TCGGTAGGTACACAGGAGTTGAAGGAGGCGATGTGTTTCATATGCGCAGCGTCATC 1536
QY 1599 CAGTAGCATCAACCGCTCTCCGAACTCCATCTCTGCTGGCTTCGGTATCAACGCTGAA 1658
Db 1537 CAGTAGCATCAACCGCTCTCCGAACTCCATCTCTGCTGGCTTCGGTATCAACGCTGAA 1596
QY 1659 ACAACCAAGAAATCTTCTTGCAGGTGATAGGACAAATGTGTATAGACAGATAGAGAAGC 1718
Db 1597 ACAACCAAGAAATCTTCTTGCAGGTGATAGGACAAATGTGTATAGACAGATAGAGAAGC 1656
QY 1719 AAGCGAAGGATTTAGCATTTCCCTGGGTGGGTGAACAAAGTTGAGAAAGCTTCATCAAAAACC 1778

Db 1657 AAGCGAAGGATTTAGCATTTCCCTGGTTCCGGTGAACAAAGTTGAGAAGCTTCATCAAAAACC 1716
QY 1779 AGAAGGAATCTCATTCTGAGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCGTCTC 1838
Db 1717 AGAGGAGTCTCATTCTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTC----- 1765
QY 1839 CTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTC 1898
Db 1766 -----GTCCTCTGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTC 1815
QY 1899 CACTCTTTCAATTTGAAGCTTTTAATCTGAGAAATGAGGCAACTTGTATGTATCGAT 1958
Db 1816 CACTCTTTCAATTTGAAGCTTTTAATCTGAGAAATGAGGCAACTTGTATGTATCGAT 1875
QY 1959 AATAGATCAGCTTTTGTACTTACTATCAAAATCTTCAATAAATAAAGGTTTG 2018
Db 1876 AATAGATCAGCTTTTGTATCTACTATCAAAATCTTCAATAAATAAAGGTTTG 1935
QY 2019 TGCCTTTGTTCTCC 2032
Db 1936 TGCCTTTGTTCTCC 1949

RESULT 12
AF432231
LOCUS
DEFINITION
Arachis hypogaea major allergen Arahi gene, complete cds.
ACCESSION
AF432231
VERSION
AF432231.1
KEYWORDS
GI:16612199
SOURCE
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE
1 (bases 1 to 2674)
Viquez, O.M., Konan, K.N. and Dodo, H.W.
Structure and organization of the genomic clone of a major peanut
allergen gene, Ara h 1
Mol. Immunol. 40 (9), 565-571 (2003)
JOURNAL
MEDLINE
PUBMED
14597159
REFERENCE
2 (bases 1 to 2674)
Viquez, O.M., Konan, K.N. and Dodo, H.W.
Direct Submission
AUTHORS
Submitted (11-OCT-2001) Food and Animal Sciences, Alabama A&M
University, 4900 Meridian Street North, Normal, AL 35762, USA
JOURNAL
TITLES
Location/Qualifiers
1. 2674
/organism="Arachis hypogaea"
/molecule="genomic DNA"
/variety="F78-1339"
/db_xref="taxon:3818"
join(<1. 811,883. 1058,1308. 1388,1463. 2674)
/product="major allergen Arahi"
<1. 811
/number=1
<1. 90
join(91. 811,883. 1058,1308. 1388,1463. 2365)
/codon_start=1
/product="major allergen Arahi"
/protein_id="AAL27476.1"
/db_xref="GI:16612200"
/translation="MRGRVSPFLMLLLGLVLASVSATHAKSPYOKKTENPCARCLQ
SCQEPDLKQKACESTKLEYDPRCVYDPRGHTGTTNQSPKIRPERGEDEWGT
DRQPREEGRWGPAGEREREDWRPFSHQPRKIRPERGEDEWGT
PGSHVRETSNNPFYFRRFSTRYQNGRIRVLORFQDORSQFONLQNHRI VOIE
AKPNTLVLPKHADADNILLVQOQATVTVANGNNRKSFLNDEGHALRIPSGFISYLN
RHDNQLRVAKISNPVTPGQFEPFPPASSRDSQSLQGFSENTLEAFNFEIR
VLLSEAGGEQGRKSTRSENNEGVIVKVSKEHVELTKHAKSVSKGSEEG
DITNPILNREGEPLDNNFGKLFVKPKKPNQLQDLDMMLTCTVEIKEGALMLPHFNS

exon	883..1058	KAMVIVVVKGTGKGNLELVAVRKEQQQRREDEDEEEEGSGNRVRRVYARLKEGD
	/number=2	
exon	1308..1388	VFTMPAHPVAIIINASSLHLGFGINAENRIFLAGDKDNVIDQIEKQADLAFPGS
	/number=3	
exon	1463..2674	GEQVEKLIIKNQKESHFVSAR PQSQSPSPSPKESPEKQEENQGGKPLLSILKA
	/number=4	
3' UTR	2366..2674	
polyA_signal	2440..2445	
ORIGIN		

Query Match	79.1%;	Score 1608;	DB 8;	Length 2674;
Best Local Similarity	83.8%;	Pred. No. 0;		
Matches 2032;	Conservative 0;	Mismatches 0;	Indels 394;	Gaps 3;
Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGACGAGGACGAATGAGAGGGAG	60	
Db	42	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGACGAGGACGAATGAGAGGGAG	101	
Qy	61	GGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTCGCTTCAGTTTCTGCAAGCCA	120	
Db	102	GGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTCGCTTCAGTTTCTGCAAGCCA	161	
Qy	121	TGCCAAGTCATCACTTTACAGAGAAGAAAACAGAGAACCCCTGCGCCACAGAGGTGCTTCCA	180	
Db	162	TGCCAAGTCATCACTTTACAGAGAAGAAAACAGAGAACCCCTGCGCCACAGAGGTGCTTCCA	221	
Qy	181	GAGTTGTCAACGGAACCGGATGACTTGAAGCAAAAGGCAATGCGAGTCTCGCTGCACCAA	240	
Db	222	GAGTTGTCAACGGAACCGGATGACTTGAAGCAAAAGGCAATGCGAGTCTCGCTGCACCAA	281	
Qy	241	GCTCGAGTATGATCTCTGTTGTGTCTATGATCTCTCGAGGACACACTGCGACCAACCA	300	
Db	282	GCTCGAGTATGATCTCTGTTGTGTCTATGATCTCTCGAGGACACACTGCGACCAACCA	341	
Qy	301	ACGTTCCCTCCAGGGAGCGGACACGTGCGCGCCAAACCGCGAGACTPACGATGATGACCG	360	
Db	342	ACGTTCCCTCCAGGGAGCGGACACGTGCGCGCCAAACCGCGAGACTPACGATGATGACCG	401	
Qy	361	CCGTCAACCCCGAAGAGAGGAGGCCGATGCGGACCAAGTTGGAACCGAGGAGCGTGA	420	
Db	402	CCGTCAACCCCGAAGAGAGGAGGCCGATGCGGACCAAGTTGGAACCGAGGAGCGTGA	461	
Qy	421	AAGAGAGAGACA CTGGAGACAA CCAAGAGAAGATTGGAGCGGACCAAGTCATCAGCAGCC	480	
Db	462	AAGAGAAGAAGA CTGGAGACA CAA CCAAGAGAAGATTGGAGCGGACCAAGTCATCAGCAGCC	521	
Qy	481	ACGGAAAAATAGGCCCGAAGAGAGAGAAAGAGTGGGGAAACACACAGGTAGCCCA	540	
Db	522	ACGGAAAAATAGGCCCGAAGAGAGAGAAAGAGTGGGGAAACACACAGGTAGCCCA	581	
Qy	541	TGTGAGGGAAAGAAACATCTCGGAACAAACCTTCTACTTCCGTCAAGCGGTTTAGCAC	600	
Db	582	TGTGAGGGAAAGAAACATCTCGGAACAAACCTTCTACTTCCGTCAAGCGGTTTAGCAC	641	
Qy	601	CCGCTACGGGAAACCAAAACGGTAGGATCCGGGTCTCTGACAGGTTTGACAAAGGTCAG	660	
Db	642	CCGCTACGGGAAACCAAAACGGTAGGATCCGGGTCTCTGACAGGTTTGACAAAGGTCAG	701	
Qy	661	GCAGTTTCAGAACTCCAGAAATCAACCGTATTTGTGCAGATCGAGGCCAAACCTTAACACTCT	720	
Db	702	GCAGTTTCAGAACTCCAGAAATCAACCGTATTTGTGCAGATCGAGGCCAAACCTTAACACTCT	761	
Qy	721	TGTTCTTCCCAAGCAGCTGATCGTGAATAACCTCTTGTTATCCAGCA	768	
Db	762	TGTTCTTCCCAAGCAGCTGATCGTGAATAACCTCTTGTTATCCAGCAAGGATCAAAATC	821	
Qy	769	-----A	769	
Db	822	TAAATCTCTTAAACTACATATATTTGTGTCTTGATACATATGATTCATTCGATTCGA	881	

QY	770	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAG	829
DB	882	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAAATCTTTGACGAG	941
QY	830	GGCCATGCACCTCAGAAATCCCAATCCGGTTTCAATTTCTTACATCTTTGAAACCGGCATGACAAC	889
DB	942	GGCCATGCACCTCAGAAATCCCAATCCGGTTTCAATTTCTTACATCTTTGAAACCGGCATGACAAC	1001
QY	890	CAGAACTCTCAGAGTAGCTTAAAAATCTCAATGCCCGTTAAACACACCCCGCCAGTTTG	944
DB	1002	CAGAACTCTCAGAGTAGCTTAAAAATCTCAATGCCCGTTAAACACACCCCGCCAGTTTGAGGTA	1061
QY	945	-----	944
DB	1062	CCTCTTCTTTTACATATATATTCAGAGGAATGTTAGGAGCCAGCAACTTTTGTAAAT	1121
QY	945	-----	944
DB	1122	TTGTAGTTATCAAAATAGTCATTAGTGATGATATTAATGTTTATTAAATAATGATTTTAAATGG	1181
QY	945	-----	944
DB	1182	TGTAATAATTTTATCTAATGGTTCAATTTTTTTTTTTTGATATCATGTTGGTCAGAAATTTAAC	1241
QY	945	-----	944
DB	1242	AAATTAAGTGGTCTAGATTTTTTCTTATATTCAAATTTCAATTTATCATCTTTACATGTTGCT	1301
QY	945	-----AGAAATTTCTCCGGGAGCAGCGGAGACCAATCATCTCTACTTTGCAAGGGCTTCAGC	1000
DB	1302	TCACAGGATTTCTTTCCGGGAGCAGCGGAGACCAATCATCTCTACTTTGCAAGGGCTTCAGC	1361
QY	1001	AGGAATACGTTGGAGGCGGCTTCAAT-----	1027
DB	1362	AGGAATACGTTGGAGGCGGCTTCAATGTAAGCAATGTATCATTAATTAGGAATTTAAAT	1421
QY	1028	-----GCGGAATTTCAATGAGATAC	1046
DB	1422	AACGTATATGTTATAATAAACTTATAATATATACATAGGCGGAATTTCAATGAGATAC	1481
QY	1047	GGAGGTGCTGTTAGAGAGNAATGCAGGAGGTGACGAAGGAGGAGAGGCGCAGAGCGCAT	1106
DB	1482	GGAGGTGCTGTTAGAGAGNAATGCAGGAGGTGACGAAGGAGGAGAGGCGCAGAGCGCAT	1541
QY	1107	GGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGCAGC	1166
DB	1342	GGAGTACTCGGAGTAGTGAGNAACATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGCAGC	1601
QY	1167	TTGAAGAACTTTAATAAGCAAGCTTAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAG	1226
DB	1602	TTGAAGAACTTTAATAAGCAAGCTTAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAG	1661
QY	1227	ATATCCAAACCCCAATCAAATTCGAGAAAGCGCAGCCCGATCTTTCTTCAAACAATTTCCGGA	1286
DB	1662	ATATCCAAACCCCAATCAAATTCGAGAAAGCGCAGCCCGATCTTTCTTCAAACAATTTCCGGA	1721
QY	1287	AGTTATTTTGGTGAAGCCAGACAAGAAAGACCCCGAGCTTCAGGACCTCGACATGATGC	1346
DB	1722	AGTTATTTTGGTGAAGCCAGACAAGAAAGACCCCGAGCTTCAGGACCTCGACATGATGC	1781
QY	1347	TCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCCACTTCAAATCTCAAAGGCCA	1406
DB	1782	TCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCCACTTCAAATCTCAAAGGCCA	1841
QY	1407	TGGTTATCGTCGTCTCAACAAAGGAACCTTGAACTCGTGGCTGTGAAGAAAG	1466
DB	1842	TGGTTATCGTCGTCTCAACAAAGGAACCTTGAACTCGTGGCTGTGAAGAAAG	1901
QY	1467	AGCAACAAACAGAGGGAGCGGGGAGAGAGAGGAGCGAAGACGAAGAGAGAGAGGGAA	1526
DB	1902	AGCAACAAACAGAGGGAGCGGGGAGAGAGAGGAGCGAAGACGAAGAGAGAGAGGGAA	1961

QY 1527 GTAACAGAGAGGTGGCTAGGTACACAGCAGAGGTGAAGGAGCGATGTTTCATCATGC 1586
Db GTAAACAGAGAGGTGGCTAGGTACACAGCAGAGGTGAAGGAGCGATGTTTCATCATGC 2021
QY 1587 CAGCAGCTCATCAGTAGGACCATCAAGCGTTCCTCGAATCCTCATCTGCTTGGCTTCGGTA 1646
Db CAGCAGCTCATCAGTAGGACCATCAAGCGTTCCTCGAATCCTCATCTGCTTGGCTTCGGTA 2081
QY 1647 TCAACGCTCAAAACACCAAGAGTCTTCTTGCAGGTGATAGGACATGATGATAGACC 1706
Db TCAACGCTCAAAACACCAAGAGTCTTCTTGCAGGTGATAGGACATGATGATAGACC 2141
QY 1707 AGATAGAGAGCAAGCAAGAGTATAGCATTCCTCGGTCGGGTGAACAAAGTTGAGAAGC 1766
Db AGATAGAGAGCAAGCAAGAGTATAGCATTCCTCGGTCGGGTGAACAAAGTTGAGAAGC 2201
QY 1767 TCATCAAAACACCAAGAGTATCTCTTTGTGAGTGTCTGCTCAATCTCAATCTCAAT 1826
Db TCATCAAAACACCAAGAGTATCTCTTTGTGAGTGTCTGCTCAATCTCAATCTCAAT 2261
QY 1827 CTCGTCGTCCTCTGAGAAGAGTCTCTGAGAAGAGTCAAGAGGAGGAAACCAAG 1886
Db CTCGTCGTCCTCTGAGAAGAGTCTCTGAGAAGAGTCAAGAGGAGGAAACCAAG 2321
QY 1887 GAGGAAGGTCCTCTCTTTCAATTTTGAAGGCTTTAACTGAGAAATGGAGGCAACTTG 1946
Db GAGGAAGGTCCTCTCTTTCAATTTTGAAGGCTTTAACTGAGAAATGGAGGCAACTTG 2381
QY 1947 TTATGATCGATAATAGATCAGCTTTTGTGATCTTACTATCCAAATTTATCAATAA 2006
Db TTATGATCGATAATAGATCAGCTTTTGTGATCTTACTATCCAAATTTATCAATAA 2441
QY 2007 TAAAAACGTTCTGCGTCTGTTCTCC 2032
Db TAAAAACGTTCTGCGTCTGTTCTCC 2467

RESULT 13
AY581852
LOCUS AY581852 2332 bp DNA linear PLN 01-MAY-2004
DEFINITION Arachis hypogaea conarachin gene, partial cds.
ACCESSION AY581852
VERSION AY581852.1 GI:46560477
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 2332)
Li,H., Wang,L., Liao,B., Yan,Y., Lin,X. and Huang,S.
Molecular cloning of conarachin gene from peanut seed
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 2332)
Li,H., Wang,L., Liao,B., Yan,Y., Lin,X. and Huang,S.
Direct Submission
TITLE
Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
GuangDong 510275, China
Location/Qualifiers
FEATURES
source
1..2332
/organism="Arachis hypogaea"
/mol_type="genomic DNA"
/cultivar="Sanfou523"
/db_xref="taxon:3818"
Join(343..1045,1117..1292,1369..1449,1523..>2303)
/product="conarachin"
Join(343..1045,1117..1292,1369..1449,1523..>2303)
/codon_start=1
/product="conarachin"
/protein_id="AA00597.1"
/db_xref="GI:46560478"

/translation="MRGRVSPIMLLILVLASVATQAKSPYRKTNPCARQCIQSC
QOEPPDLKQKACSRCTKLEYPQVDTGATNQRHPGRTGQPDQYDDRRQPR
REEGMDGAPREREREEDWDOPREDWRRDSHOQPRKIRPEGREGEWEQWTPGSEVR
BETSRNPPYPSPRFSRYNQNGRIRVLQRFQSRKQFQNLQNRHIVQLEARNPIL
VLKHADNINLVIQQGAIVTVANGNNRKSFNLDGHALRIPSGFISYILNRHQN
LRVAKISNPNTPGQFFPASSRDSYLOFSRNTLEAFAFNFEISYILNRHJEN
AGGQBERGQRRTSRSDNKGVIKVSKEHVELTHAKSVSKSGSEEDITNPIL
RDGPDLNPNFGRLLFEVKPDKNLQLOLDMLTCEIKEGALMLPHNSKAMVTVV
NKGTVNLELVAVRKEOQGRREOEWEEREEREESNRVRRVTARLKEGDVIMP
AAHPVNASSELHLLGFGINAENHRIPLAGDKDNVIDQIEKQAKDLAFPGSGQVE
KLINQRSHFVSA"

ORIGIN

Query Match 68.7%; Score 1395.4; DB 8; Length 2332;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 46; Indels 253; Gaps 8;

QY 2 ATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAGG 61
Db ATAATAATACATATATTCATCAATCATCTATATATAAGTAGTTGCAGGAGCAATGAGAGGAGG 354
QY 62 GTTTCCTCACTGATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCAT 121
Db GTTTCCTCACTGATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCAG 414
QY 122 GCCAAGTCATCACTTACAGAGAAAGAAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCAG 181
Db GCCAAGTCACCC-----TTACGGGNAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCAG 468
QY 182 AGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCTCGTGTGCAACCAAG 241
Db AGTTGTCAACAGGAACCGGACGACTTGAAGCAAAAGGATCGAGTCTCGTGTGCAACCAAG 528
QY 242 CTCGAGTATGATCTCTGTTGTCTATGATCCTCGAGGACACACTGCGCACCAACCAACCA 301
Db CTCGAGTATGATCTCTGTTGTCTATG-----ACACTGGCGCCACCAACCAACCA 576
QY 302 CGTTCCCTCCAGGGGAGCGACACGTCGCGCCCAACCCGGAGACTAGCATGATGACCGC 361
Db CGTCACCTCCAGGGGAGCGACACGTCGCGCCCAACCCGGAGACTAGCATGATGACCGC 636
QY 362 CGTCAACCCCGAAGAGAGGAGCGCGATGGGGAACAGCTGGACCGAGGAGGCGTGAA 421
Db CGTCAACCCCGAAGAGAGGAGCGCGATGGGGAACAGCTGAAACCGAGGAGGCGTGAA 696
QY 422 AGAGNAGAGACTGAGACACCAACAGAGAAAGTTGGAGCGACCAAGTCATCAGCAGCA 481
Db AGAGNAGAGACTGAGACACCAACAGAGAAAGTTGGAGCGACCAAGTCATCAGCAGCA 756
QY 482 CGGAAATTAAGGCCCGAAGGAGAGAGAGCAAGAGTGGGGAAACCAAGGTAGCCAT 541
Db CGGAAATTAAGGCCCGAAGGAGAGAGAGCAAGAGTGGGGAAACCAAGGTAGCCAG 816
QY 542 GTGAGGGAAGAAACATCTCGGAAACAAACCTTTCTATCTTCCTCGTCAAGCGGTTTAGCACC 601
Db GTGAGGGAAGAGACATCACGGAACAAACCTTTCTACTTCTCCGTCAAGCGGTTTAGCACC 876
QY 602 CGCTACGGGAACCAAAACGGTAGGATCGGGTCCTCGAGAGGTTTGACCAAGGTCRAGG 661
Db CGCTACGGGAACCAAAACGGTAGGATCGGGTCCTCGAGAGGTTTGACCAAGGTCRAGG 936
QY 662 CAGTTTCAGAAATCTCCAGAATCACCGTATTTGTGAGATCGAGGCGCAACCACTCTT 721
Db CAGTTTCAGAAATCTCCAGAATCACCGTATTTGTGAGATCGAGGCGCAACCACTCTT 996
QY 722 GTTCTTCCCAAGCAGCTGTGATGAACATCTCTGTTTATCCAGCA----- 768
Db GTTCTTCCCAAGCAGCTGTGATGAACATCTCTGTTTATCCAGCAAGGATCAAAATCT 1056
QY 769 -----AG 770
Db AATCTTACTTACTATATATATTTTGTGTTGATACATATGATTCATGGATTGAG 1116

QY	820	TCTTGACGAGGCGCATGCACTCAGAACTCCATCCGGTTCCTTCAATTCCTCATCTTGAACCG	879
Db	182	TCTTGACGAGGCGCATGCACTCAGAACTCCATCCGGTTCCTTCAATTCCTCATCTTGAACCG	241
QY	880	CCATGACACACGAACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGGGCA	939
Db	242	CCATGACAAACGAACTCAGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGGGCA	301
QY	940	GTTTGAGGATTTCTTCCCGCGCAGACGAGACCAATCATCTTCTTCAGGGCTTCAG	999
Db	302	GTTTGAGGATTTCTTCCCGCGCAGACGAGACCAATCATCTTCTTCAGGGCTTCAG	361
QY	1000	CAGGAATACGTTGGAGGCCCTTCAATCGGAATTCATAGATACGAGGGTCTGTT	1059
Db	362	CAGGAATACGTTGGAGGCCCTTCAATCGGAATTCATAGATACGAGGGTCTGTT	421
QY	1060	AGAAGAGATCAGAGGTCAGCAAGAGAGAGAGGGCGAGCGATGAGTACTCGGAG	1119
Db	422	AGAAGAGATCAGAGGTCAGCAAGAGAGAGAGGGCGAGCGATGAGTACTCGGAG	481
QY	1120	TAGTGAGAACAAATGAAGAGTGTAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTAC	1179
Db	482	TAGTGAGAACAAATGAAGAGTGTAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTAC	541
QY	1180	TAAGCACGCTAAATCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACC	1239
Db	542	TAAGCACGCTAAATCGTCTCAAAGAAAGGCTCC--GAAGAGGAGATATCACCAACC	598
QY	1240	AATCAACTTGAGAGAGGCGCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGT	1299
Db	599	AATCAACTTGAGAGAGGCGCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGT	658
QY	1300	GAAGCAGACAAAGAAACCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGA	1359
Db	659	GAAGCAGACAAAGAAACCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGA	718
QY	1360	GATCAAGAGAGGCTTTGATGCTCCACATTCACATCAAAGGCGATGTTATCGTGT	1419
Db	719	GATCAAGAGAGGCTTTGATGCTCCACATTCACATCAAAGGCGATGTTATCGTGT	778
QY	1420	CCTCAACAAAGAACTGGAAACTTGAACTCTGGCTGTAGAAAAGAGCAACACAGAG	1479
Db	779	CCTCAACAAAGAACTGGAAACTTGAACTCTGGCTGTAGAAAAGAGCAACACAGAG	838
QY	1480	GGGACGGCGGAAGAGAGGAGCAAGACAAAGAGAGGAGGAAGTAAACAGAGAGGT	1539
Db	839	GGGACGGCGGAAGAGAGGAGCAAGACAAAGAGAGGAGGAAGTAAACAGAGAGGT	898
QY	1540	GGGTAGGTACACAGCGAGGTTGAAGGAAGGGCATGTGTTTCATCATGCCAGCAGCTCATCC	1599
Db	899	GGGTAGGTACACAGCGAGGTTGAAGGAAGGGCATGTGTTTCATCATGCCAGCAGCTCATCC	958
QY	1600	AGTAGCCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAA	1659
Db	959	AGTAGCCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAA	1018
QY	1660	CAACCAACAGAACTTTCTTGACAGGTATGAAGACAAATGTATAGACCAAGAGAGCA	1719
Db	1019	CAACCAACAGAACTTTCTTGACAGGTATGAAGACAAATGTATAGACCAAGAGAGCA	1078
QY	1720	ACGGAAGGATTTAGATTCCTCGGTCGGGTGAAACAAAGTTGAGAGCTCATCAAAAAACA	1779
Db	1079	ACGGAAGGATTTAGATTCCTCGGTCGGGTGAAACAAAGTTGAGAGCTCATCAAAAAACA	1138
QY	1780	GAAGAACTCACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCC	1839
Db	1139	GAAGAACTCACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCC	1198
QY	1840	TGAGAAAGAGTCTCTTGAGAAAGAGATCAAGAGGAGGAAAACCAAGGAGGAGGGTCC	1899
Db	1199	TGAGAAAGAGTCTCTTGAGAAAGAGATCAAGAGGAGGAAAACCAAGGAGGAGGGTCC	1258
QY	1900	ACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATA	1959

Db	1259	ACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATA	1318
QY	1960	ATAAGATCAGCTTTGTACTCTACTATCCAAAACTTATCAATAAATAAAACGTTTGT	2019
Db	1319	ATAAGATCAGCTTTGTACTCTACTATCCAAAACTTATCAATAAATAAAACGTTTGT	1378
QY	2020	GGTTGTTTCTCC	2032
Db	1379	GGTTGTTTCTCC	1391
RESULT 15			
AY581850			
LOCUS	AY581850	1035 bp	mRNA linear PLN 01-MAY-2004
DEFINITION	Arachis hypogaea conarachin mRNA, partial cds.		
ACCESSION	AY581850		
VERSION	AY581850.1	GI:46560473	
KEYWORDS	Arachis hypogaea (peanut)		
SOURCE	Arachis hypogaea		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.		
REFERENCE	Wang, L., Liao, B., Li, H., Yan, Y., Lin, X. and Huang, S.		
AUTHORS	1 (bases 1 to 1035)		
TITLE	cDNA cloning of conarachin in peanut seed		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1035)		
AUTHORS	Wang, L., Liao, B., Li, H., Yan, Y., Lin, X. and Huang, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAR-2004) Department of Biology and Biotechnology, School of Life Science, Sun Yat-sen University, Guangzhou, Guangdong 510275, China		
FEATURES	Location/Qualifiers		
source	1..1035		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/cultivar="SanYou523"		
	/db_xref="taxon:3818"		
	<1..901		
	/note="similar to peanut allergen 1 (p17)"		
	/codon_start=2		
	/product="conarachin"		
	/protein_id="AA00595.1"		
	/translation="LEAFNAEFNEIRRVLEENAGGEQERGGRRWSTRSENNEGIVKVKSHVEBLTKHAKSVSKGSEEGDITNPINLRGEPDLNNPKLFVQPKDK NPOQLDMLTCEIKEGALMLPHFNSKAMVIVVKNKGTNLNVAVRKQQRGR EEEDEQDEBGSNRVRYTARLKEGDFVTPAAHPVAINASSELHLGFGINAENN HRIPLAGDKNDVIDIEKQAKDLAPGSGEQVEKLKNQRESHFVSARPOQSPPSE KEDDEENQGGKGLLSILKAFN"		
CDS			
ORIGIN			
Query Match	47.0%	Score 954.4;	DB 8; Length 1035;
Best Local Similarity	96.9%	Pred. No. 2.3e-258;	
Matches	992;	Conservative 0;	Mismatches 11; Indels 21; Gaps 1;
QY	1009	GTTGAGGCCCCCTTCAATCGGAATTCATAGATACGAGGGTCTGTTAGAGAGAA	1068
Db	1	GTTGAGGCCCCCTTCAATCGGAATTCATAGATACGAGGGTCTGTTAGAGAGAA	60
QY	1069	TGCAGGAGTGCAGCAAGAGGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAAA	1128
Db	61	TGCAGGAGTGCAGCAAGAGGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAAA	120
QY	1129	CAATCAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCACGTTCAAGAACTTACTAAGCAGC	1188
Db	121	CAATCAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCACGTTCAAGAACTTACTAAGCAGC	180
QY	1189	TAATCCGTCTCAAAGAAAGGCTCCGAAGAAAGAGGAGATATCACCAACCAACTT	1248

Db	181	TAAATCCGCTCTCAAGAAAGGCTCCGAAAGAGGGAGATATACCAACCCCAATCAACTT	240
Qy	1249	GAGAGAAGCGGAGCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGA	1308
Db	241	GAGAGAAGCGGAGCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGA	300
Qy	1309	CAAGAGAAGACCCCGACGCTTCAGAGCCTGGACATGATCTCACCTGTGTAGAGATCAAGA	1368
Db	301	CAAGAAGAAGACCCCGACGCTTCAGAGCCTGGACATGATCTCACCTGTGTAGAGATCAAGA	360
Qy	1369	AGGAGCTTTGATGCTCCCACTTCAACTCAAAGGCCATGGTTATCTGTCGTCTCAACAA	1428
Db	361	AGGAGCTTTGATGCTCCCACTTCAACTCAAAGGCCATGGTTATCTGTCGTCTCAACAA	420
Qy	1429	AGGAACCTGGAACCTTGAACTCGTGTCTGTAGAAAAGCAACAACAGAGGGGACGGCG	1488
Db	421	AGGAACCTGGAACCTTGAACTCGTGTCTGTAGAAAAGCAACAACAGAGGGGACGGCG	480
Qy	1489	GGAGAAGAGGAGGACGAAGAGCAAGCAAGAGAGGGAAGTAACAGAGAGGTGCTAGGTA	1548
Db	481	GGAGAAGAGGAGGACGAAGAGCAAGCAAGAGAGGGAAGTAACAGAGAGGTGCTAGGTA	540
Qy	1549	CACAGCGAGGTTGAAGGAAGGCGATGTGTTCAATCATGCCAGCAGCTCATCCAGTAGCCAT	1608
Db	541	CACAGCGAGGTTGAAGGAAGGCGATGTGTTCAATCATGCCAGCAGCTCATCCAGTAGCCAT	600
Qy	1609	CAACGCTTCTCCGAATCTCACTGCTGTGGCTTCGGTATCAACGCTGAAAAACAACACAG	1668
Db	601	CAACGCTTCTCCGAATCTCACTGCTGTGGCTTCGGTATCAACGCTGAAAAACAACACAG	660
Qy	1669	AATCTTCTTGCAGGTGATAGGACAACTGTGATAGCAGATAGAGAAACAGCGGAGGA	1728
Db	661	AATCTTCTTGCAGGTGATAGGACAACTGTGATAGCAGATAGAGAAACAGCGGAGGA	720
Qy	1729	TTTAGCATTTCCCTGGGTCGGGTGAACAAAGTTGAGAAAGCTCATCAAAAACAGAAAGGAATC	1788
Db	721	TTTAGCATTTCCCTGGGTCGGGTGAACAAAGTTGAGAAAGCTCATCAAAAACAGAAAGGAATC	780
Qy	1789	TCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGA	1848
Db	781	TCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCGTCTCTCTGAGAAAGA	819
Qy	1849	GTCTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGGAAGGGTCCACTCCTTTC	1908
Db	820	GTCTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGGAAGGGTCCACTCCTTTC	879
Qy	1909	AATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAAGATCA	1968
Db	880	AATTTTGAAGGCTTTTAACTGAGAAATGGAGGCAACTTGTATGTATCGATATAAGATCA	939
Qy	1969	CGCTTTTGTACTTACTATCCAAAACTTATCAATAAATAAAAAAGTTTGTGCGTTGTTT	2028
Db	940	CGCTTTTGTAACTACTATCCAAAACTTATCAATAAATAAAAAAGTTTGTGCGTTGTTT	999
Qy	2029	CTCC 2032	
Db	1000	CTCC 1003	

Search completed: August 24, 2005, 00:57:06

Job time : 6030.11 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 4765.9 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataacatatattcatc.....cgttgtggtgttcttc 2032

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gses1.*
9: gb_gses2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	719	35.4	721	6	CD038628
2	713	35.1	714	6	CD038277
3	711.4	35.0	723	6	CD038527
4	681	33.5	717	6	CD038837
5	657.2	32.3	724	6	CD038648
6	629.2	31.0	719	6	CD038172
7	626	30.8	676	6	CD038790
8	602	29.6	685	6	CD038555
9	576	28.3	588	6	CD038119
10	527.6	26.0	684	6	CD038540
11	473	23.3	474	6	CD038694
12	471	23.2	582	6	CD038620
13	458.8	22.5	509	6	CD038253
14	436	21.5	484	6	CD038101
15	428.6	21.1	443	7	CO897502
16	425	20.9	425	6	CD038284
17	419.6	20.6	536	6	CD038075
18	392	19.3	411	7	CO897503
19	358.6	17.6	406	6	CD038765
20	307.6	15.1	823	6	CA858229
21	306.2	15.1	851	6	CA858769
22	296.8	14.6	894	6	CA858664
23	296.6	14.6	806	4	BI310422
24	292.2	14.4	864	5	BQ123471

25	290.8	14.3	780	4	BI311149
26	290.6	14.3	823	5	BQ122514
27	290	14.3	862	5	BQ123135
28	285.8	14.1	854	5	BQ123411
29	285.4	14.0	830	6	CA858151
30	283	13.9	805	4	BI312339
31	281.4	13.8	760	6	CA858137
32	281.4	13.8	769	4	BI311022
33	280.2	13.8	847	6	CA858293
34	278.4	13.7	781	4	BI311613
35	274.6	13.5	765	5	BQ122510
36	273.2	13.4	804	6	CA858167
37	272.4	13.4	843	4	BI311012
38	271.8	13.4	805	6	CA858433
39	271.4	13.4	734	6	CA858790
40	271.4	13.4	839	4	BI311235
41	271.2	13.3	861	6	CA858864
42	271	13.3	860	5	BQ122453
43	270.4	13.3	701	5	BQ124840
44	270.4	13.3	727	4	BI309850
45	270.4	13.3	837	4	BI310945

ALIGNMENTS

RESULT 1
LOCUS CD038628 721 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI010_A05 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI010_A05 5', mRNA sequence.
ACCESSION CD038628
VERSION CD038628.1 GI:30420466
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 721)
AUTHORS Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G. and Lynch, R.E.
TITLE Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL Unpublished (2003)
COMMENT Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES

source Location/Qualifiers
1..721
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI010_A05"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCVILXAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent and ultrapurified (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN		Query Match	35.4%;	Score 719;	DB 6;	Length 721;	
		Best Local Similarity	99.7%;	Pred. No. 4.3e-191;			
		Matches 719;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1231	CACCAACCCCAATCAACTTGAGAGAGCGCGAGCCGATCTTTCTAACAACTTTGGGAAGTT	1290				
Db	1	CACCAACCCCAATCAACTTGAGAGAGCGCGAGCCGATCTTTCTAACAACTTTGGGAAGTT	60				
QY	1291	ATTGTAGGTGAAGCCACAGACAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCAC	1350				
Db	61	ATTGTAGGTGAAGCCACAGACAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCAC	120				
QY	1351	CTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCCATGGT	1410				
Db	121	CTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCCATGGT	180				
QY	1411	TATCGTCGTCGTCACAAAGGAACCTGGAAACCTTGAACTCGTGGCTGTGAAGAAAGAGCA	1470				
Db	181	TATCGTCGTCGTCACAAAGGAACCTGGAAACCTTGAACTCGTGGCTGTGAAGAAAGAGCA	240				
QY	1471	ACACAGAGGGGACGGCGGAAGAGGAGGAGCGAAGACGAGAGAGAGAGGAGGAGTAA	1530				
Db	241	ACACAGAGGGGACGGCGGAAGAGGAGGAGCGAAGACGAGAGAGAGAGGAGGAGTAA	300				
QY	1531	CAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTATCATGCGCAGC	1590				
Db	301	CAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTATCATGCGCAGC	360				
QY	1591	AGCTCATCCAGTAGCCATCAACCGCTTCCCTCCGAACTCCCATCTGCTTGGCTTCGGTATCAA	1650				
Db	361	AGCTCATCCAGTAGCCATCAACCGCTTCCCTCCGAACTCCCATCTGCTTGGCTTCGGTATCAA	420				
QY	1651	CGCTGAAACCAACACAGATCTTCCCTGCGAGGTATAGGACAAATGTGATAGACCAGAT	1710				
Db	421	CGCTGAAACCAACACAGATCTTCCCTGCGAGGTATAGGACAAATGTGATAGACCAGAT	480				
QY	1711	AGAGAAGCAAGCGAAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAAGTTGAGAAGCTCAT	1770				
Db	481	AGAGAAGCAAGCGAAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAAGTTGAGAAGCTCAT	540				
QY	1771	CAAAAACCGAAGGAATCTCATTTGTGTAGTGTCTGCTCTCAATCTCAATCTCAATCTCC	1830				
Db	541	CAAAAACCGAAGGAATCTCATTTGTGTAGTGTCTGCTCTCAATCTCAATCTCAATCTCC	600				
QY	1831	GTGCTCTCTGACAAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAAACCAAGGAGS	1890				
Db	601	GTGCTCTCTGACAAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAAACCAAGGAGS	660				
QY	1891	GAAGGGTCCACTCTTTTCAATTTTGAAGGCTTTTAACTGAGAATGGAGGCAACTTCTTAT	1950				
Db	661	GAAGGGTCCACTCTTTTCAATTTTGAAGGCTTTTAACTGAGAATGGAGGCAACTTCTTAT	720				
QY	1951	G 1951					
Db	721	G 721					
RESULT 2		CD038277	714 bp	mRNA	linear	EST 07-MAY-2003	
LOCUS		CD038277					

UTPPI005 A03 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI005_A03 5', mRNA sequence.
CD038277
CD038277.1 GI:30420115
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 714)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1. .714
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI005_A03"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80OC freezer.
Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN		Query Match	35.1%;	Score 713;	DB 6;	Length 714;	
		Best Local Similarity	99.9%;	Pred. No. 2.1e-189;			
		Matches 713;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	146	AAAAACAGAGAACCCCTCGGCCAGAGGTCCTCCAGAGTTGTCAACAGGAACCGGATGAC	205				
Db	1	AAAAACAGAGAACCCCTCGGCCAGAGGTCCTCCAGAGTTGTCAACAGGAACCGGATGAC	60				
QY	206	TTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTC	265				
Db	61	TTGAAGCAAAAGGCATGCGAGTCTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTGTC	120				
QY	266	TATGATCTCTGAGGACACACTGCGCACCAACCAACCTCCCTCCAGGAGCGGACCA	325				
Db	121	TATGATCTCTGAGGACACACTGCGCACCAACCAACCTCCCTCCAGGAGCGGACCA	180				

```
Qy 326 CGTGGCCGCAACCCGAGACTACGATGATGACCGCGTCAACCCCGAAGAGAGGA 385
Db 181 CGTGGCCGCAACCCGAGACTACGATGATGACCGCGTCAACCCCGAAGAGAGGA 240
Qy 386 GGCCGATGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAGAGAGAGAGAGGA 445
Db 241 GGCCGATGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAGAGAGAGAGGA 300
Qy 446 AGAGAAGATTGGAGGCGACCAAGTCATCAGCAGCCAGGAAATTAAGCCCGAAGGA 505
Db 301 AGAGAAGATTGGAGGCGACCAAGTCATCAGCAGCCAGGAAATTAAGCCCGAAGGA 360
Qy 506 GAAGGAGAACAGAGTGGGAGACCAAGTCATCAGCAGCCAGGAAATTAAGCCCGAAG 565
Db 361 GAAGGAGAACAGAGTGGGAGACCAAGTCATCAGCAGCCAGGAAATTAAGCCCGAAG 420
Qy 566 AACCTTTTCTACTTCCCGTCAAGGCGGTTTACGACCCGCTACGGGAAACCAAAACGGTAGG 625
Db 421 AACCTTTTCTACTTCCCGTCAAGGCGGTTTACGACCCGCTACGGGAAACCAAAACGGTAGG 480
Qy 626 ATCCGGTCTCTGAGAGTTTACCAAAAGTCAAGGAGTTTCAAGATCTCCAGATCAC 685
Db 481 ATCCGGTCTCTGAGAGTTTACCAAAAGTCAAGGAGTTTCAAGATCTCCAGATCAC 540
Qy 686 CGTATTGTGAGATCGAGGCGCAACCTTAACACTCTTCTTCCCAAGCAGCTGATGCT 745
Db 541 CGTATTGTGAGATCGAGGCGCAACCTTAACACTCTTCTTCCCAAGCAGCTGATGCT 600
Qy 746 GATAACATCTTGTATTCAGCAGGCGCAACCTTACCGTAGCCGTAAGCAATGGCAATAAC 805
Db 601 GATAACATCTTGTATTCAGCAGGCGCAACCTTACCGTAGCCGTAAGCAATGGCAATAAC 660
Qy 806 AGAAGAGCTTTAATCTTACGAGGCGCATGCACTCAGATCCCATCCGTTTC 859
Db 661 AGAAGAGCTTTAATCTTACGAGGCGCATGCACTCAGATCCCATCCGTTTC 714

RESULT 3
CD038527 723 bp mRNA linear EST 07-MAY-2003
LOCUS UTPPI008_E12 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION Arachis hypogaea cDNA clone UTPPI008_E12 5', mRNA sequence.
ACCESSION CD038527
VERSION CD038527.1 GI:30420365
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 723)
Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
and Lynch, R. E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..723
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI008_E12"
/tissue_type="Immature pods"
/dev_stage="R6"

FEATURES
source
```

```
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV11XA4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

ORIGIN

Query Match	35.0%	Score 711.4	DB 6	Length 723
Best Local Similarity	99.4%	Pred. No. 6e-189		
Matches 712	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy 1210	CTCCGAAGAGAGGAGATATCACCAACCCAACTCAACTTGCAGAGAGGCGAGCCGATCT	1269		
Db 1	CTCCGAAGAGAGGAGATATCACCAACCCAACTCAACTTGCAGAGAGGCGAGCCGATCT	60		
Qy 1270	TTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGCAAGAAACCCCGAGTTCA	1329		
Db 61	TTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGCAAGAAACCCCGAGTTCA	120		
Qy 1330	GGACCTGGACATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCACA	1389		
Db 121	GGACCTGGACATGCTCACCTGTGTAGAGATCAAGAGGAGCTTTGATGCTCCACA	180		
Qy 1390	CTTCAACTCAAGGCCATGTTATCGTCGTCGTCACAAAGGAACTGGAACCTTGAAC	1449		
Db 181	CTTCAACTCAAGGCCATGTTATCGTCGTCGTCACAAAGGAACTGGAACCTTGAAC	240		
Qy 1450	CGTGGCTGTAAGAAAAGAGCAACAAGAGGGGACCGCGGAAGAGAGGAGGACGAAGA	1509		
Db 241	CGTGGCTGTAAGAAAAGAGCAACAAGAGGGGACCGCGGAAGAGAGGAGGACGAAGA	300		
Qy 1510	CGAAGAGAGGAGGAGTAACAGAGAGGTCGTAGGTACACAGCGAGGTTGAGGAGG	1569		
Db 301	CGAAGAGAGGAGGAGTAACAGAGAGGTCGTAGGTACACAGCGAGGTTGAGGAGG	360		
Qy 1570	CGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCATCAACGCTTCCTCCGAACCTCCA	1629		
Db 361	CGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCATCAACGCTTCCTCCGAACCTCCA	420		
Qy 1630	TCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATAA	1689		
Db 421	TCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATAA	480		
Qy 1690	GGCAATGTGATAGACAGATAGAGAGCAGAGGATTTAGCATTCCTCGGTCGGG	1749		
Db 481	GGCAATGTGATAGACAGATAGAGAGCAGAGGATTTAGCATTCCTCGGTCGGG	540		
Qy 1750	TGAACAAGTTGAGAAGCTCATCAAAAAACAGAAAGGAATCTCACTTTGTGAGTCTCGTCC	1809		
Db 541	TGAACAAGTTGAGAAGCTCATCAAAAAACAGAAAGGAATCTCACTTTGTGAGTCTCGTCC	600		
Qy 1810	TCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCA	1869		
Db 601	TCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCA	660		
Qy 1870	AGAGGAGGAAAACCAAGGAGGAGGGTCCACTCTCTTCAATTTTGNAGGCTTTTA	1925		

```
|||||
661 AGAGGAGAAACCGAGGGGAGGGTGCTCACTCTTCNATTTTGAAGGCTTTTA 716

RESULT 4
CD038837
LOCUS
DEFINITION
UTPPI012 H04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI012_H04 5', mRNA sequence.
ACCESSION
CD038837
VERSION
CD038837.1 GI:30420675
KEYWORDS
EST.
SOURCE
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 717)
REFERENCE
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..717
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTPPI012_H04"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar Al3
(NCV11XR4). Al3 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise phagescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match 33.5%; Score 681; DB 6; Length 717;
Best Local Similarity 99.6%; Pred. No. 2.3e-180;
Matches 714; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 967 CCGAGACCAATCATCTACTTGCAGGGCTTCACAGGAATACGTTGGAGCCGCTTCAA 1026
|||||
Db 1 CCGAGACCAATCATCTACTTGCAGGGCTTCACAGGAATACGTTGGAGCCGCTTCAA 60
|||||
QY 1027 TCGGGAATTCATGAGATACGGAGGGTGCTGTTTGAAGAATGTCAGGAGGTGAGCAAGA 1086
|||||
```

```
Db 61 TCGGGAATTCATGAGATACGGAGGGTGCTGTTTGAAGAATGTCAGGAGGTGAGCAAGA 120
|||||
QY 1087 GGAGAGGGCGAGAGCGGATGCGAGTACTCGGAGTACTGAGACAAATGAAGGAGTGTAGT 1146
|||||
Db 121 GGAGAGAGGGCGAGAGCGGATGCGAGTACTCGGAGTACTGAGACAAATGAAGGAGTGTAGT 180
|||||
QY 1147 CAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTCAAAGAA 1206
|||||
Db 181 CAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTCAAAGAA 240
|||||
QY 1207 AGGCTCCGGAAGAAGGAGGAGATATCAACCAATCAACTTGAAGAGGAGGAGGCCGA 1266
|||||
Db 241 AGGCTCCGGAAGAAGGAGGAGATATCAACCAATCAACTTGAAGAGGAGGAGGCCGA 300
|||||
QY 1267 TCTTTTCAACRACTTTGGGAAGTATTATGAGGTGAAGCCAGACAAAGAAACCCCGAGCT 1326
|||||
Db 301 TCTTTTCAACRACTTTGGGAAGTATTATGAGGTGAAGCCAGACAAAGAAACCCCGAGCT 360
|||||
QY 1327 TCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCC 1386
|||||
Db 361 TCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCC 420
|||||
QY 1387 ACACCTCAACTCAAAGGCCATCGTTATCGTCTGTCAAAGAGAACTTGAACCTTGA 1446
|||||
Db 421 ACACCTCAACTCAAAGGCCATCGTTATCGTCTGTCAAAGAGAACTTGAACCTTGA 480
|||||
QY 1447 ACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGCGCGGAGAGAGAGGAGCA 1506
|||||
Db 481 ACTCGTGGCTGTAAAGAAAGACCAACAGAGGGGCGCGGAGAGAGAGGAGCA 540
|||||
QY 1507 AGACGAAGAAG-AGGAGGGAAGTAACAGA-GAGGTGCGGTAGGTACACACGCGAGTTGAAG 1564
|||||
Db 541 AGACGAAGAAGAGGAGGGAAGTAACAGAGGAGGTGCGGTAGGTACACACGCGAGTTGAAG 600
|||||
QY 1565 GAAGCGGATGTTTATCATATGCCAGAGCTATCAGTAGTCCATCAACGCTTCTCTCCGA 1624
|||||
Db 601 GAAGCGGATGTTTATCATATGCCAGAGCTATCAGTAGTCCATCAACGCTTCTCTCCGA 660
|||||
QY 1625 CTCATCTGTTGGCTTCGGTATCAACGCTGAAAAACAACACAGATCTTCTTGTGA 1681
|||||
Db 661 CTCATCTGTTGGCTTCGGTATCAACGCTG-AAACACCAACAGATCTTCTTGTGA 716

RESULT 5
CD038648
LOCUS
DEFINITION
UTPPI010 C11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI010_C11 5', mRNA sequence.
ACCESSION
CD038648
VERSION
CD038648.1 GI:30420486
KEYWORDS
EST.
SOURCE
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 724)
REFERENCE
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..724
FEATURES
source
```

```
/organism="Arachis hypogaea"
/mol_type="mRNA"
/db_xref="taxon:3818"
/clone="UTPPI010_C11"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar Al3 (NCV11XAR4). Al3 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise phagescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

ORIGIN

Query Match 32.3%; Score 657.2; DB 6; Length 724;
Best Local Similarity 95.9%; Pred. No. 1.2e-173;
Matches 697; Conservative 0; Mismatches 24; Indels 6; Gaps 2;

Qy 799 CAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACTCAGATCCCATCCGGTTT 858
Db 1 CAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACTCAGATCCCATCCGGTTT 60

Qy 859 CATTTCTACATCTTGAAACCGCCATGACCAACCAAGCTTCAGAGTAGCTAAAATCTCCAT 918
Db 61 CATTTCTACATCTTGAAACCGCCATGACCAACCAAGCTTCAGAGTAGCTAAAATCTCCAT 120

Qy 919 GCCGGTTAACACCCGGCCAGTTTGAGGATTTCTTCGGCGGAGGAGCGGAGACCAATC 978
Db 121 GCCGGTTAACACCCGGCCAGTTTGAGGATTTCTTCGGCGGAGGAGCGGAGACCAATC 180

Qy 979 ATCTACTTCGAGGCTTCAGCAGGAATAGCTTGGAGCGCGCTTCAATCGGAATTCAA 1038
Db 181 ATCTACTTCGAGGATTCAGCAGGAATATTTGGAGCGCGCTTCAATCGGAATTCAA 240

Qy 1039 TGAGATACGAGGGTGTGTTAGAAAGAAATGCAAGGAGGTGAGCAAGAGGAGAGGGCA 1098
Db 241 TGAGATACGAGGGTGTGTTAGAAAGAAATGCAAGGAGGTGAGCAAGAGGAGAGGGCA 300

Qy 1099 GAGCGATGAGTAGTACTCGGAGTAGTGAAACAATGAAGGAGTAGTAGTCAAAGTGCAMA 1158
Db 301 GAGCGCAGGAGTACTCGGAGTAGTG---ATAATGAAGGAGTAGTAGTCAAAGTGCAMA 357

Qy 1159 GGAGCAGTTGAAGAACTTACTAGCAGCGTAAATCCGTCCTCAAGAAAGCGTCCGAAGA 1218
Db 358 GGAGCAGTTCAAGAACTTACTAGCAGCGTAAATCCGTCCTCAAGAAAGCGTCC---GA 414

Qy 1219 AGAGGAGATATCACAACCAATCAACTTGAGAGAGGGGAGCGCCGATCTTTCTAACAA 1278
Db 415 AGAGGAGATATCACAACCAATCAACTTGAGAGAGGGGAGCGCCGATCTTTCTAACAA 474

Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAGAAACCCAGCTTCAGGACCTGGA 1338
Db 475 CTTTGGGAGGTTATTTGAGGTGAAGCCAGACAGAGAAACCCAGCTTCAGGACCTGGA 534

Qy 1339 CATGATGCTCACTGTGTAGAGATCAAGAGGAGCTTTTGTGCTCCACACTCAACTC 1398
Db 1398 CATGATGCTCACTGTGTAGAGATCAAGAGGAGCTTTTGTGCTCCACACTCAACTC 1398

Db 535 CATGATGCTCACTGTGTAGAGATCAAGAGGAGCTTTTGTGCTCCACACTCAACTC 594

Qy 1399 AAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGAAACCTTGAACCTCGGCTGT 1458
Db 595 AAAGGCCATGTTATCGTCTGCTCAACAAAGAACTGGAAACCTTGAACCTCGTGTGT 654

Qy 1459 AAGAAAGACCAACAGAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518
Db 655 AAGAAAGAGCNACACAGAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 714

Qy 1519 GGAGGGA 1525
Db 715 TGAAGAA 721

RESULT 6

CD038172

LOCUS

DEFINITION

CD038172 719 bp mRNA linear EST 07-MAY-2003

Arachis hypogaea cDNA clone UTPPI003_D10 5', mRNA sequence.

CD038172

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arachis hypogaea (peanut)

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 719)

and Lynch, R.E.

Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

CONTACT: Baozhu Guo

Molecular Genetics

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3

Location/Qualifiers

1. .719

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="Al3"

/db_xref="taxon:3818"

/clone="UTPPI003_D10"

/tissue_type="Immature pods"

/dev_stage="R6"

/lab_host="XLI-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar Al3 (NCV11XAR4). Al3 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise phagescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

QY 1194 CCGTCTCAAGAGAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTTGAGAG 1253
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 361 CCGTCTCAAGAGAGGCTCCGAAGAAGAGGAGATATACCAACCCCAATCAACTTGAGAG 420
 QY 1254 AAGCGAGCGCCGATCTTTCTAAACAACCTTTGGGAAGTTATTGTGAGTGAAGCCAGCAAG- 1312
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 421 AAGCGAGCGCCGATCTTTCTAAACAACCTTTGGGAAGTTATTGTGAGTGAAGCCAGCAAGN 480
 QY 1313 AAGAACCCCAAGCTTCAGGACCTGGACATGATGCTCA-CTGTGTAGAGATCAAGAAGG 1371
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 481 AAGAACCCCAAGCTTCAGGACCTGGACATGATGCTCACTCCCTGTGTAGAGATCAAGAAGG 540
 QY 1372 AGCTTTGATGCTCCACACTTCACATCAAGGCGCATGTTATCTGCTGCTCAACAAGG 1431
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 541 AGCTTTGATGCTCCACACTTCACATCAAGGCGCATGTTATCTGCTGCTCAACAAGG 600
 QY 1432 AACTTGGAAACCTTGAACTCTGCTGGCTGTAAAGAAAGAGCAACAACAGAGGGGACGGCGGG- 1490
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 601 AACTTGGAAACCTTGAACTCTGCTGGCTGT-AGNAAAGAGCAACACNAGAGGGGACGGCGGGA 659
 QY 1491 AAGAAGAGGAGGAGCA 1507
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 660 AAGAAGAGGAGGAGCA 676

RESULT 8

CD038555

LOCUS

DEFINITION

Arachis hypogaea cDNA clone UTPPI009_A05 5', mRNA sequence.

ACCESSION

CD038555

KEYWORDS

EST.

SOURCE

Arachis hypogaea (peanut)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;

Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 685)

AUTHORS

Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.

and Lynch, R. E.

TITLE

Generation and Analyses of ESTs for Arachis hypogaea

JOURNAL

Unpublished (2003)

COMMENT

Contact: Baozhu Guo

Molecular Genetics

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd, Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

FEATURES

source

1..685

Location/Qualifiers

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPPI009_A05"

/tissue_type="Immature pods"

/dev_stage="R6"

/lab_host="X1-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;

cDNA library was constructed from peanut cultivar A13;

(NCV1XAR). A13 has resistance to Aspergillus infection

and drought tolerance. The immature pods that developed to

R6 stage were collected from different plants, and placed

into liquid N2 immediately and stored in -80oC freezer.

Total RNA was isolated with TRIzol-Reagent

ultrapure (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 29.6%; Score 602; DB 6; Length 685;
 Best Local Similarity 94.3%; Pred. No. 4.5e-158;
 Matches 663; Conservative 0; Mismatches 21; Indels 19; Gaps 3;

QY 29 ATATAAGTAGTAGCAGGAGCAATGAGAGGAGGGTTCCTCCACTGATGCTGTGCTAGGG 88
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 1 ATATAAGTAGTAGCAGGAGCAATGAGAGGAGGGTTCCTCCACTGATGCTGTGCTAGGG 60
 QY 89 ATCTTTGTCCTGGCTTCAGTTTCTGCAACGCAATGCCAAGTCATCACCTTACCAGAA 148
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 61 ATCTTTGTCCTGGCTTCAGTTTCTGCAACGCAATGCCAAGTCACCTTACCAGAA 114
 QY 149 ACAGAGAACCCCTGCGCCAGAGGTGCTTCAGAGTTGTCAACAGGAAACCGGATGACTTG 208
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 115 ACAGAGAAACCCCTGCGCCAGAGGTGCTTCAGAGTTGTCAACAGGAAACCGGATGACTTG 174
 QY 209 AAGCAAAAGCATGCGAGTCTGCTCCACCAAGCTCGAGTATGATCTCTGCTGTGCTAT 268
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 175 AAGCAAAAGCATGCGAGTCTGCTCCACCAAGCTCGAGTATGATCTCTGCTGTGCTAT 234
 QY 269 GATCTCTGAGGACACACTGCGACCAACCAACCAAGCTTCCCTCCAGGGAGCGGACACGT 328
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 235 G-----ACACTGGGGCCACCAACCAAGCTCACCTCCAGGGAGCGGACACGT 282
 QY 329 GCGCGCAACCCGAGACTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 388
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 283 GCGCGCAACCCGAGACTACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 QY 389 CGATGGGACCACTGCGACCGGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 343 CGATGGGACCACTGCGACCGGAGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
 QY 449 GAAGATTGGAGCGCACCAAGTTCATCAGCAGCCACCGAAATAAGGCCCGGAGGAGAGAGAA 508
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 403 GAAGATTGGAGCGCACCAAGTTCATCAGCAGCCACCGAAATAAGGCCCGGAGGAGAGAGAA 462
 QY 509 GGAGAACAGAGTGGGGAAACACAGGTAGCCATGTGAGGAGAGAGAGAGAGAGAGAGAGAG 567
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 463 GGAGAACAGAGTGGGGAAACACAGGTAGCCATGTGAGGAGAGAGAGAGAGAGAGAGAGAG 522
 QY 568 CCCTTTCTACTTCCCTCAAGCGGTTTACGACCCCTAGCGGAGAGAGAGAGAGAGAGAGAGAG 627
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 523 CCCTTTCTACTTCCCTCAAGCGGTTTACGACCCCTAGCGGAGAGAGAGAGAGAGAGAGAGAG 582
 QY 628 CCGGGTCTGAGAGGTTTGACCAAGGTCAAGCGAGTTTTCAGAAATCTCCAGAGATCACCG 687
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 583 CCGGTCTGAGAGGTTTGACCAAGGTCAAGCGAGTTTTCAGAAATCTCCAGAGATCACCG 642
 QY 688 TATTGTGAGATGAGGCGCAACCTAACACTCTTGTGTTCTTCCC 730
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 643 TATTGTGAGATGAGGCGCAGACCTAACACTCTTGTGTTCTTCCC 685

RESULT 9

CD038119

LOCUS

DEFINITION

Arachis hypogaea cDNA clone UTPPI002_F11 5', mRNA sequence.

ACCESSION

CD038119

VERSION

CD038119.1 GI:30419957

CD038119 588 bp mRNA linear EST 07-MAY-2003
 UTPPI002_F11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
 Arachis hypogaea cDNA clone UTPPI002_F11 5', mRNA sequence.

construction followed the protocol of by ZAP-cDNA Gigapack IIII Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI/EcoRI sites adapters. The lambda library was packed into phages using Gigapack IIII Gold (Stratagene). The un-amplified library was used to excise phagescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match	26.0%;	Score 527.6;	DB 6;	Length 684;
Best Local Similarity	94.3%;	Pred. No. 4.5e-137;		
Matches 584;	Conservative 0;	Mismatches 20;	Indels 15;	Gaps 3;
Qy	957	CGGCGACGCCGAGACCAATCATCTTACCTTGACGGGCTTCAGACAGAAATACGTTGAGG	1016	
Db	72	CGGCGACGCCGAGACCAATCATCTTACCTTGACGGGATTCAGCAGAAATCTTTTGAGG	131	
Qy	1017	CCGCTTCAATTCGGAATTCAATCAGATACGAGGGTGCTTTAGAAGAGAAATCAGAG	1076	
Db	132	CCGCTTCAATTCGGAATTCAATCAGATACGAGGGTGCTTTAGAAGAGAAATCAGAG	191	
Qy	1077	GTGAGCAAGAGGAGAGGGCAGAGGCGATGGAGTACTCGAGTAGTGAAGAAACAATGAAG	1136	
Db	192	GAGAGCAAGAGGAGAGGGCAGAGGCGAGCGAGTACTCGAGTAGTGTG--ATAATCAAG	248	
Qy	1137	GAGTGATAGTCAAAAGTGTCAAAGAGGACGTTTGAAGAACTTACTAAGCAGCGCTAAATCCG	1196	
Db	249	GAGTGATAGTCAAAAGTGTCAAAGAGGACGTTTGAAGAACTTACTAAGCAGCGCTAAATCCG	308	
Qy	1197	TCTCAAGAAAGGCTCCGAGAGAGAGGGAGATATCACCAACCCCAATCAACTTGTGAGAGAG	1256	
Db	309	TCTCAAGAAAGGCTCC--GAAGAGAAAGATATCACCAACCCCAATCAACTTGTGAGAGATG	365	
Qy	1257	GCAGGCCGATCTTTCTAAACAATTTTCGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAGA	1316	
Db	366	GCAGGCCGATCTTTCTAAACAATTTTCGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAGA	425	
Qy	1317	ACCCCCAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTT	1376	
Db	426	ACCCCCAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTT	485	
Qy	1377	TGATGCTCCACACTTCAACTCAAGAGCCATGGTTATCGTGTGCTGTCTCAACAAAGGAACTG	1436	
Db	486	TGATGCTCCACACTTCAACTCAAGAGCCATGGTTCATCGTGTGCTGTCTCAACAAAGGAACTG	545	
Qy	1437	GAACCTTGAATCTCGTGGCTGTGAAGAAAGACAAACAAGAGGGGACGCGGAA---1492		
Db	546	GAACCTTGAATCTCGTGGCTGTGAAGAAAGACAAACAAGAGGGGACGCGGAAACAAG	605	
Qy	1493	-----GAAGAGGAGGACGAAGACGAAGAGGAGGGAAGTAAACAGAGAGGTGCGTAGGT	1547	
Db	606	AGTCGGAAGAAGAGGAGGAAGATGAAGAAAGAGGAGGAGTAAACAGAGAGGTGCGTAGGT	665	
Qy	1548	ACACAGCGAGGTTGAAGGA	1566	
Db	666	ACACAGCGAGGTTGANGAA	684	

RESULT	11
LOCUS	CD038694
DEFINITION	UTPP1010_H08 USDA-Riflon Peanut mRNA linear EST 07-MAY-2003 Arachis hypogaea cDNA clone UTPP1010_H08 5', mRNA sequence.
ACCESSION	CD038694
VERSION	CD038694.1
KEYWORDS	EST.
SOURCE	GI:30420532
ORGANISM	Arachis hypogaea (peanut)
	Arachis hypogaea
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurooids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
1 (bases 1 to 474)
Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bao, and Lynch, R. E.
Generation and analyses of ESTs for *Arachis hypogaea* Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seg primer: T3.

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
1. J. H.
2. J. H.
3. J. H.
4. J. H.
5. J. H.
6. J. H.
7. J. H.
8. J. H.
9. J. H.
10. J. H.
11. J. H.
12. J. H.
13. J. H.
14. J. H.
15. J. H.
16. J. H.
17. J. H.
18. J. H.
19. J. H.
20. J. H.
21. J. H.
22. J. H.
23. J. H.
24. J. H.
25. J. H.
26. J. H.
27. J. H.
28. J. H.
29. J. H.
30. J. H.
31. J. H.
32. J. H.
33. J. H.
34. J. H.
35. J. H.
36. J. H.
37. J. H.
38. J. H.
39. J. H.
40. J. H.
41. J. H.
42. J. H.
43. J. H.
44. J. H.
45. J. H.
46. J. H.
47. J. H.
48. J. H.
49. J. H.
50. J. H.
51. J. H.
52. J. H.
53. J. H.
54. J. H.
55. J. H.
56. J. H.
57. J. H.
58. J. H.
59. J. H.
60. J. H.
61. J. H.
62. J. H.
63. J. H.
64. J. H.
65. J. H.
66. J. H.
67. J. H.
68. J. H.
69. J. H.
70. J. H.
71. J. H.
72. J. H.
73. J. H.
74. J. H.
75. J. H.
76. J. H.
77. J. H.
78. J. H.
79. J. H.
80. J. H.
81. J. H.
82. J. H.
83. J. H.
84. J. H.
85. J. H.
86. J. H.
87. J. H.
88. J. H.
89. J. H.
90. J. H.
91. J. H.
92. J. H.
93. J. H.
94. J. H.
95. J. H.
96. J. H.
97. J. H.
98. J. H.
99. J. H.
100. J. H.

FEATURES
SOURCE

```

Location/Qualifiers
1. .474
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTPP1010_H08"
/tissue_type="Immature pods"
/dev_host="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
(UTPP)"
/note="vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar Al3
(NC811X4R4). Al3 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIZOL-Reagent
ultrapure(GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

```

ORIGIN

Query Match	23.3%	Score 473;	DB 6;	Length 474;
Best Local Similarity	99.8%;	Pred. No. 1e-121;		
Matches 473;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1559	TTCAAGGAGGGCGATGTGTTTCATCATGCCGACGAGCTCATCCAGCTAGCCATCAACGGCTTCC	1618	
Db	1	TTGAAGGAAAGGCGATGTGTTTCATCATGCCGACGAGCTCATCCAGTAGGCATCAACGGCTTCC	60	
Qy	1619	TCCGAACTCCATCTGCTTGGCTTCGGGTATCAACGCTTGAAAAACAAACACAGAACTCTTCCTT	1678	
Db	61	TCCGAACTCCATCTGCTTGGCTTCGGGTATCAACGCTTGAAAAACAAACACAGAACTCTTCCTT	120	
Qy	1679	GCAGGTGATAAGGACAAATGTGATAGACCATAGAGAGCAAGCAAGGATTTAGCAATTC	1738	
Db	121	GCAGGTGATAAGGACAAATGTGATAGACCATAGAGAGCAAGCAAGGATTTAGCAATTC	180	
Qy	1739	CCTGGGTCGGGTGAACAAAGTTGAGAGAGCTCATCAAAAAACAGAAAGNAATCTCACTTTGTG	1798	
Db	181	CCTGGGTCGGGTGAACAAAGTTGAGAGAGCTCATCAAAAAACAGAAAGNAATCTCACTTTGTG	240	
Qy	1799	AGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTCAGAAAGAGTCTCCTTGAG	1858	
Db	241	AGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTCAGAAAGAGTCTCCTTGAG	300	
Qy	1859	AAAGAGGATCAAGGAGGAGAAAAACCAAGGAGGGAAGGGTCCACTCTCTTTCAATTTTGAAG	1918	

Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1..509
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI004_F03"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="Xl1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XA4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fractionation. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 22.6%; Score 458.8; DB 6; Length 509;
Best Local Similarity 93.8%; Pred. No. 1.1e-11;
Matches 497; Conservative 0; Mismatches 12; Indels 21; Gaps 1;

QY 1494 AGACGAGCAGCAAGAACGAAGAGCGGGAGTAAACAGAGAGTGCGTAGGTACACAG 1553
DB 1 AAGACAGAGGAGNATGAAGAAGAGGAGGAGTAACAGAGAGTGCGTAGGTACACAG 60

QY 1554 CGAGGTTGAAGGAAGCGCATGTGTTTCATCATGCCAGCGCTCATCCAGTACCATCAACG 1613
DB 61 CGAGGTTGAAGGAAGCGCATGTGTTTCATCATGCCAGCGCTCATCCAGTACCATCAACG 120

QY 1614 CTTCTCCGAATCTCAATCTGCTTGGCTTCGGTATCAAGCTGMAAACACACAGAAATCT 1673
DB 121 CTTCTCCGAATCTCAATCTGCTTGGCTTCGGTATCAAGCTGMAAACACACAGAAATCT 180

QY 1674 TCCTTCAGGTATAGGACAATGTATAGCACAGATAGAGCAAGCAAGGATTAG 1733
DB 181 TCCTTCAGGTATAGGACAATGTATAGCACAGATAGAGCAAGGATTAG 240

QY 1734 CATTCCTCGGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAAAGAAATCTCACT 1793
DB 241 CATTCCTCGGTCGGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAGTCTCACT 300

QY 1794 TTGTGAGTGTCTGCTTCAATCTCAATCTCGGTTCGCTCTCTGAGAAAGAGTCTC 1853
DB 301 TTGTGAGTGTCTGCTTCAATCTCAATCTCGGTTCGCTCTCTGAGAAAGAGTCTC 339

QY 1854 CTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCTCTCTCAATTT 1913
DB 340 CTGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCTCTCTCAATTT 399

QY 1914 TGAAGGCTTTTAACTGAGAAATCGAGCAACTGTTTATGTATCGATAATAAGATCACGCTT 1973
DB 400 TGAAGGCTTTTAACTGAGAAATCGAGCAACTGTTTATGTATCGATAATAAGATCACGCTT 459

QY 1974 TTGTACTCTACTATCAAATACTTATCAATAAATAAAACGTTTGTGCGT 2023

Db 460 TTGTAATCTACTATCAAATACTTATCAATAAATAAAACGTTTGTGCGT 509

RESULT 14
CD038101
LOCUS
DEFINITION
UTPPI002_E01 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI002_E01 5', mRNA sequence.
ACCESSION
CD038101
VERSION
CD038101.1 GI:30419939
KEYWORDS
EST
SOURCE
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 484)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES
Location/Qualifiers
1..484
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI002_E01"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="Xl1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XA4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fractionation. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 21.5%; Score 436; DB 6; Length 484;
Best Local Similarity 93.8%; Pred. No. 2.9e-11;
Matches 473; Conservative 0; Mismatches 10; Indels 21; Gaps 1;

QY 1505 GAAGACGAAGAGAGGAGGAGTAACAGAGAGTGCGCTAGGTACACAGCGAGTTGAAG 1564
DB 2 GAAGATGAAGAGAGGAGGAGTAACAGAGAGTGCGCTAGGTACACAGCGAGTTGAAG 61

QY 1565 GAAGCGATGTCTTCATCATGCCAGCACTCATCCAGTAGCCATCAACGCCTCTCTCCGAA 1624

```
Db 62 GAAGGGGATGTGTTTCATGCGCAGCGCTCATCCAGTAGCCATCAACGGTCTCCTCCGAA 121
QY 1625 CTCATCTGCTTGGCTTCGGTATCAAGCTGAAACAAACACACAGAATCTTCTTGCAGGT 1684
Db 122 CTCATCTGCTTGGCTTCGGTATCAAGCTGAAACAAACACAGAATCTTCTTGCAGGT 181
QY 1685 GATAAGGACAATGTGTAGACCAAGTAGAGAGCAAGCGAAGATTTAGCATTTCCCTGGG 1744
Db 182 GATAAGGACAATGTGTAGACCAAGTAGAGAGCAAGCGAAGATTTAGCATTTCCCTGGT 241
QY 1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGAATCTCACTTTGTGAGTGCT 1804
Db 242 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGGAGTCTCACTTTGTGAGTGCT 301
QY 1805 CGTCCTCAATCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTGAGAAAGAG 1864
Db 302 CGTCCTCAATCTCAATCTCGGTC-----GTCCTCTGAAAAAGAG 340
QY 1865 GATCAAGAGGAGGAAAAACCAAGGAGGAGGGTCCACTCTCTTTCAATTTTGAAGGCTTTT 1924
Db 341 GATCAAGAGGAGGAAAAACCAAGGAGGAGGGTCCACTCTTTCAATTTTGAAGGCTTTT 400
QY 1925 AACTGAGATGGAGGCAACTGTTGATGATCGATAATAAGATCACGCTTTTGTACTCTAC 1984
Db 401 AACTGAGATGGAGGAACTGTTGATGATGCCATAATAAGATCACGCTTTTGTAACTAC 460
QY 1985 TATCCAAAACCTTATCAATAAATA 2008
Db 461 TATCCAAAACCTTATCAATAAATA 484
```

```
RESULT 15
LOCUS CO897502
DEFINITION CO897502 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.
ACCESSION CO897502
VERSION CO897502.1 GI:51237292
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 443)
Yan, Y.S., Wang, L. and Huang, S.Z.
Expressed sequence tags of mid-matured cotyledons of peanut
Unpublished (2004)
Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (860) 02084036592
Email: yys9803@yahoo.com.cn
Seq primer: pTriplex2 Forward.
FEATURES
source
1..443
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/tissue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/notes="Organ: Seed; Vector: lambdaTriplex2"
```

ORIGIN

```
Query Match 21.1%; Score 428.6; DB 7; Length 443;
Best Local Similarity 98.0%; Pred. No. 3.4e-109;
Matches 434; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 597 GCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAAGT 656
Db 1 GCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGACCAAAGT 60
QY 657 CAAAGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCGAGATCGAGGCCAAACCTTAA 716
Db 61 CAAAGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCGAGATCGAGGCCAAACCTTAA 120
QY 717 CTCCTTGTCTTCCCAAGCAGCGTGTGATAAATCCTTGTATCCAGCAAGGSCAAG 776
Db 121 CTCCTTGTCTTCCCAAGCAGCGTGTGATAAATCCTTGTATCCAGCAAGGSCAAG 180
QY 777 CCACCGTGACCGTAGCAAAATGCAATAACAGAAAAGAGCTTTAATCTTGAAGGGCCATG 836
Db 181 CCACCGTGACCGTAGCAAAATGCAATAACAGAAAAGAGCTTTAATCTTGAAGGGCCATG 240
QY 837 CACTCAGAAATCCATCCGGTTTCATTTCTTGAACCGCCCATGTAACCCAGAAC 896
Db 241 CACTCAGAAATCCATCCGGTTTCATTTCTTGAATCGATGTAACCCAGAAC 300
QY 897 TCAGAGTAGCTAAAATCTCCATGCCGTTTAAACACCCCGCCAGTTTGAGGATTTCTTCC 956
Db 301 TCAGAGTAGCTAAAATCTCCATGCCGTTTAAACACCCCGCCAGTTTGAGGATTTCTTCC 360
QY 957 CGCGCAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTTCAGCAGGAATACGTTGGAG 1016
Db 361 CGCGCAGCAGCCGAGACCAATCATCTACTTGCAGGGATTCAGCAGGAATACGTTGGAG 420
QY 1017 CGGCCTTCAATCGGGAATTCAT 1039
Db 421 CGGCCTTCAATCGGGAATTCAT 443
```

Search completed: August 24, 2005, 03:44:08
Job time : 4769.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 760.871 Seconds
(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataatcatatattcatc.....cgtttgtcggtgtttttcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	4 AAF90339	Aaf90339 Peanut al
2	2032	100.0	2032	8 ABX70603	Abx70603 Peanut Ar
3	2032	100.0	2032	10 ADG27462	Adg27462 Peanut Ar
4	2032	100.0	2041	4 AAS08537	Aas08537 DNA encod
5	2025.6	99.7	2032	2 AAT76613	Aat76613 Peanut al
6	1930	95.0	1930	2 AAZ06382	Aaz06382 Peanut al
7	1752.4	86.2	1949	2 AAT76612	Aat76612 Peanut al
8	1752.4	86.2	1949	8 ABX70604	Abx70604 Peanut Ar
9	1750.8	86.2	1949	10 ADG27463	Adg27463 Peanut Ar
10	1749.8	86.1	1952	4 AAS08539	Aas08539 Anaphylac
11	408.4	20.1	1251	6 ABS55196	Abs55196 Glycine m
12	406.8	20.0	1254	9 ACC49562	Acc49562 Mature be
13	406.8	20.0	1278	9 ACC49561	Acc49561 FLAG-tag
14	406.8	20.0	1320	2 AAV17564	AAV17564 Coding se
15	406.8	20.0	1320	9 ACC49553	Acc49553 Glycine m
16	404	19.9	1632	6 ABS55198	Abs55198 Glycine m
17	404	19.9	1818	2 AAV17562	AAV17562 Coding se
18	402	19.8	1680	6 ABS55197	Abs55197 Glycine m
19	324.2	16.0	1920	2 AAV17563	AAV17563 Coding se
20	160.2	7.9	2140	2 AAV42316	AAV42316 Macadamia

21	157.4	7.7	2171	2	AAV42311	AAV42311 Macadamia
22	152.6	7.5	2171	2	AAV42310	AAV42310 Macadamia
23	138.8	6.8	1867	2	AAQ20377	AAQ20377 Sequence
24	98.8	4.9	1924	2	AAV72243	AAV72243 G. max SB
25	88.4	4.4	1580	12	ADE80950	Ades80950 Cashew nu
26	88.4	4.4	1730	12	ADN80949	Adn80949 Cotton pr
27	79.6	3.9	584	13	ACN47495	Acn47495 Cotton pr
28	79.4	3.9	2332	13	ADR63468	Adr63468 Cotton CD
29	74	3.6	1722	10	ADL18492	Adl18492 Maize glo
30	74	3.6	2003	10	ADL18494	Adl18494 Maize glo
31	66	3.2	601	13	ACN47683	Acn47683 Cotton pr
32	59.4	2.9	635	13	ACN46465	Acn46465 Cotton pr
33	57.8	2.8	2000	8	ADA71938	Ada71938 Rice gene
34	56.6	2.8	574	13	ACN46319	Acn46319 Cotton pr
35	53	2.6	2010	10	ADC08559	Adc08559 Rice DNA
36	51.2	2.5	659	13	ACN50255	Acn50255 Mouse gen
37	50.2	2.5	193853	11	ACN44956	Acn44956 Mouse gen
38	49.6	2.4	10732	3	AAA10594	Aaa10594 Gene enco
39	49.4	2.4	1068	4	AAD05672	Aad05672 Human zms
40	49.4	2.4	1068	10	ADB49430	Adb49430 Degenerat
41	48.8	2.4	640	13	ADR63469	Adr63469 Cotton CD
42	47.6	2.3	591	13	ACN50453	Acn50453 Cotton ma
43	46.6	2.3	559	13	ACN48185	Acn48185 Cotton pr
44	46.6	2.3	90401	12	ADQ97515	Adq97515 Human can
45	46	2.3	564	13	ACN49247	Acn49247 Cotton pr

ALIGNMENTS

RESULT 1
AAF90339
ID AAF90339 standard; cDNA; 2032 BP.
XX
XX AAF90339;
AC
AC 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX
DE Peanut allergen Ara h1 P41B cDNA.
XX
KW Peanut; allergen; Ara h1 P41B; transgenic plant; allergy; ss.
XX
OS Arachis hypogaea.
XX
FH Key Location/Qualifiers
FT misc_feature 676..930
FT /*tag= a
FT /note= "PCR amplified fragment"
XX
FN WO200136621-A2.
XX
PD 25-MAY-2001.
XX
XX 20-NOV-2000; 2000WO-US031657.
XX
PR 19-NOV-1999; 99US-0167255P.
XX
PA (UYAL-) UNIV ALABAMA A & M.
XX
PI Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX
XX WPI; 2001-355630/37.
XX
XX Producing transgenic peanut plants that produce allergen-free seeds,
XX useful in non-allergenic foods, by antisense or sense co-suppression of
XX allergen-encoding genes.
XX
XX Claim 20; Fig 5; 72pp; English.
XX
XX The present sequence is that of peanut allergen Ara h1 P41B cDNA. A
XX portion of this gene is homologous to the corresponding region of the
XX peanut allergen Ara h1 P17 gene. This region has been PCR amplified,

1921 TTTTAACTGAGATGAGGCACTTGTATGATCAATATAGATCAAGCTTTGTACT 1980
1981 CTACTATCCAAAACCTTATCAATAAATAAAGAGTTTGCGGTGTTCTCC 2032
1981 CTACTATCCAAAACCTTATCAATAAATAAAGAGTTTGCGGTGTTCTCC 2032

RESULT 2
ABX70603
ID ABX70603 standard; cDNA; 2032 BP.
XX
AC ABX70603;
XX
DT 26-MAR-2003 (first entry)
XX
DE Peanut Ara h1 cDNA clone P41b.
XX
KW Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; IgE binding site;
KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN W0200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI; 2003-018765/01.
DR P-PSDB; ABU52412.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Example 3; Fig 10; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a cDNA
CC encoding a peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 2032 BP; 628 A; 473 C; 530 G; 401 T; 0 U; 0 Other;

Query Match 100.0%; Score 2032; DB 8; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATAATCATATATTTTCATCAATCATCTATATAAGTAGTAGCAGGCAATGAGAGGAG 60
61 GGTTCCTCCATGATGCTGTTCTAGGGATCCTTGTCTGGTTCAGTTTCTGCAACGCA 120
61 GGTTCCTCCATGATGCTGTTCTAGGGATCCTTGTCTGGTTCAGTTTCTGCAACGCA 120
121 TGCAGAGTCATCACCTTACCAGAAACAGAGAACCCCTGCGCCAGAGGTGCTCCA 180
121 TGCAGAGTCATCACCTTACCAGAAACAGAGAACCCCTGCGCCAGAGGTGCTCCA 180
181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCGCTGCAACCA 240
181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCTCGCTGCAACCA 240
241 GCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGCACCAACCA 300
241 GCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGCACCAACCA 300
301 AGTTTCCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGGAGACTACGATGATGACCG 360
301 AGTTTCCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGGAGACTACGATGATGACCG 360
361 CCGTCAACCCCGAAGAGAGGAGCGGATGGGGACCCAGCTGGACCGAGGAGGTGA 420
361 CCGTCAACCCCGAAGAGAGGAGCGGATGGGGACCCAGCTGGACCGAGGAGGTGA 420
421 AAGAGAAGAAGACTCGAGACAAACCAAGAGAAGATTGGAGGCGCAACCAAGTCAATCAGAGCC 480
421 AAGAGAAGAAGACTCGAGACAAACCAAGAGAAGATTGGAGGCGCAACCAAGTCAATCAGAGCC 480
481 ACGGAATAAGGCCCGGAGGAGAGGAGAAACAAAGAGTGGGGAACCAACCAAGGTAGCCA 540
481 ACGGAATAAGGCCCGGAGGAGAGGAGAAACAAAGAGTGGGGAACCAACCAAGGTAGCCA 540
541 TGTGAGGGAAGAAAACATCTCGGAACAAACCTTTCTACTTTCCCGTCAAGGCGGTTTAGCAC 600
541 TGTGAGGGAAGAAAACATCTCGGAACAAACCTTTCTACTTTCCCGTCAAGGCGGTTTAGCAC 600
601 CCGTACCGGGAACCAAAACCGGTAGGATCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
601 CCGTACCGGGAACCAAAACCGGTAGGATCGGGTCTCGAGAGGTTTGACCAAGGTCAAG 660
661 GCAGTTTCAGATCTCCAGAAATCAGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 720
661 GCAGTTTCAGATCTCCAGAAATCAGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 720
721 TGTTCCTCCCAAGCACGCTGATGCTGATAAATCCTTGTATCCAGCAAGGCGCAAGCCAC 780
721 TGTTCCTCCCAAGCACGCTGATGCTGATAAATCCTTGTATCCAGCAAGGCGCAAGCCAC 780
781 CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACT 840
781 CGTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACT 840
841 CAGATCCCATCCGGTTTCATTTCTCTATCTTTGAACCGCCATGACCAACCAAGCTCAG 900
841 CAGATCCCATCCGGTTTCATTTCTCTATCTTTGAACCGCCATGACCAACCAAGCTCAG 900
901 AGTAGCTAAAATCTCCATGCGCCGTTTAAACACACCCCGCCAGTTTGAAGGATTTCTTCCCGGC 960
901 AGTAGCTAAAATCTCCATGCGCCGTTTAAACACACCCCGCCAGTTTGAAGGATTTCTTCCCGGC 960
961 GAGCAGCGGAGAGGAGGCGGATGAGTACTCGGAGTGTAGTGAAGCAATGAGAGGAGTGA 1080
961 GAGCAGCGGAGAGGAGGCGGATGAGTACTCGGAGTGTAGTGAAGCAATGAGAGGAGTGA 1080
1021 CTTCAATCGGAAATTCATATGATACGAGGCGTCTGTTAGAAAGAGATGCGAGGAGTGA 1140
1021 CTTCAATCGGAAATTCATATGATACGAGGCGTCTGTTAGAAAGAGATGCGAGGAGTGA 1140
1081 GCAAGAGGAGAGGAGGCGGATGAGTACTCGGAGTGTAGTGAAGCAATGAGAGGAGTGA 1140
1081 GCAAGAGGAGAGGAGGCGGATGAGTACTCGGAGTGTAGTGAAGCAATGAGAGGAGTGA 1140

Qy	1141	GATAGTCAAAGTGTCAAAGGAGCAGCTTGAGAACTTTACTAAGCA	C	GCTTAATCCGTC	1200
Db	1141	GATAGTCAAAGTGTCAAAGGAGCAGCTTGAGAACTTTACTAAGCA	C	GCTTAATCCGTC	1200
Qy	1201	AAAGAAAGGCTCCGAAGAAGGGAGGATATCACCAACCCCAATCAA	C	CTTGAGAAAGCGA	1260
Db	1201	AAAGAAAGGCTCCGAAGAAGGGAGGATATCACCAACCCCAATCAA	C	CTTGAGAAAGCGA	1260
Qy	1261	GCCCGATCTTTCTAAACAATTTGGGAAGTTATTATTGAGTGAAG	C	CCAGCAAGAAACCC	1320
Db	1261	GCCCGATCTTTCTAAACAATTTGGGAAGTTATTATTGAGTGAAG	C	CCAGCAAGAAACCC	1320
Qy	1321	CCAGCTTCAGGA	C	CTGGACATGATGCTCACTGTGTAGAGATCAAAGAAGGAGCTTTGAT	1380
Db	1321	CCAGCTTCAGGA	C	CTGGACATGATGCTCACTGTGTAGAGATCAAAGAAGGAGCTTTGAT	1380
Qy	1381	GCTCCACACTTCAA	C	CTCRAAGGCCATGGTTATCGTCGTCGTCAA	1440
Db	1381	GCTCCACACTTCAA	C	CTCRAAGGCCATGGTTATCGTCGTCGTCAA	1440
Qy	1441	CCTTGAACTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGGGAG	C	CGCGGGAAGAAAGGA	1500
Db	1441	CCTTGAACTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGGGAG	C	CGCGGGAAGAAAGGA	1500
Qy	1501	GGACGAAGACGAAGAAGAGGGAAGTAA	C	CAGAGGTCGTTAGTGTACACAGCGAGGTT	1560
Db	1501	GGACGAAGACGAAGAAGAGGGAAGTAA	C	CAGAGGTCGTTAGTGTACACAGCGAGGTT	1560
Qy	1561	GAAGGAAGGCGATGCTGTTTCATGCCAGCAGAGCTCATCCAGT	C	AGTAGCCATCAACGCTTCC	1620
Db	1561	GAAGGAAGGCGATGCTGTTTCATGCCAGCAGAGCTCATCCAGT	C	AGTAGCCATCAACGCTTCC	1620
Qy	1621	CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGMAAAACAA	C	CCACAGAACTTCTTCC	1680
Db	1621	CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGMAAAACAA	C	CCACAGAACTTCTTCC	1680
Qy	1681	AGGTGATAGGACAAATGTGTATAGACCAAGATAGAGAAGCAAG	C	GAAGATTTAGCATTTCC	1740
Db	1681	AGGTGATAGGACAAATGTGTATAGACCAAGATAGAGAAGCAAG	C	GAAGATTTAGCATTTCC	1740
Qy	1741	TGGTCCGGGTGAACAAGTTCAGAGGCTCATCAAAACCCAGAAG	C	GAATCTCACTTTGTGAG	1800
Db	1741	TGGTCCGGGTGAACAAGTTCAGAGGCTCATCAAAACCCAGAAG	C	GAATCTCACTTTGTGAG	1800
Qy	1801	TGCTCGTCCCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAG	C	AGTCTCCTGAGAA	1860
Db	1801	TGCTCGTCCCTCAATCTCAATCTCCGTCGTCTCCTGAGAAAG	C	AGTCTCCTGAGAA	1860
Qy	1861	AGAGGATCAAGAGGAGGAAAAACCAAGAGGGAAGGGTCCACT	C	CTTTCAATTTTGAAGGC	1920
Db	1861	AGAGGATCAAGAGGAGGAAAAACCAAGAGGGAAGGGTCCACT	C	CTTTCAATTTTGAAGGC	1920
Qy	1921	TTTTTAAC	C	TGAGAAATGAGGCAACTTGTGTATGATTCGATATAAGATCA	1980
Db	1921	TTTTTAAC	C	TGAGAAATGAGGCAACTTGTGTATGATTCGATATAAGATCA	1980
Qy	1981	CTACTATCCAAAACTTTATCAATAATAAAAAAGTTTGTGCGTT	C	GTCTTCTCC	2032
Db	1981	CTACTATCCAAAACTTTATCAATAATAAAAAAGTTTGTGCGTT	C	GTCTTCTCC	2032

RESULT 3
ADG27462

ADG27462
ID ADG27462 standard; cDNA; 2032 BP.

AC ADG27462;

XX
DT 26-FEB-2004 (first entry)XX
DE Peanut Ara h1 cDNA clone P41b.

KW Peanut; plant; ss; gene; allergen; Ara h1; Ara h2; Ara h3; XX

glycinin A2B1a; Jug n1; antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound

Arachis hypogaea.

US2003202980-A1.

30-OCT-2003.

18-MAR-2002: 2002US-00100303.

29-DEC-1995: 95US-0009455P.

23-SEP-1996; 96US-00717933.
31-JAN-1998; 98US-0073283P.

13-FEB-1998; 98US-0074590P.
13-FEB-1998; 98US-0074624P.

13-FEB-1998; 98US-0074633P;
29-JUN-1998; 98US-00106872;

27-AUG-1998; 98US-00141220;
13-NOV-1998; 98US-00191593;

29-JAN-1999; 99US-00240557;
29-JAN-1999; 99US-00241101;

11-FEB-1999; 99US-00248673.
11-FEB-1999; 99US-00248674.

02-MAR-1999; 99US-0122452P.
02-MAR-1999; 99US-0122452P.

02-MAR-1999; 99US-0122360P.
02-MAR-1999; 99US-0122565P.

02-MAR-1999; 99US-0122366P;
11-MAR-1999; 99US-00267719.

Z8-JAN-2000; 2000US-00494038;
16-MAR-2001; 2001US-0276822P.

(CAPL/) CAPLAN M J.

(SOSL//) SOSIN H B.
(SAMP//) SAMPSON H.

(BANN//) BANNON G A.
(BURK//) BURKS A W.

(COCK/) COCKRELL G.
(COMP/) COMPADRE C M.

(CONN//) CONNAUGHTON C.
(HELM//) HELM R M.

(KING//) KING N E.
(KOPP//) KOPPER R A.

(PAGE//) MAHER J O.
(RABJ//) RABJOHN P A.

(STAN//) STANLEY J S.

Caplan MJ, Sosin HB, Sampso

Compadre CM, Comanagón C, Rabjohn PA, Shin DS, Stanle

WPI; 2003-875632/81.

F-FOUR, ADOZ, 403.

which has been modified so that

Continued

The invention relates to a modified ana-

anaphylactic food allergen. The natural

bond when the natural anaphylactic food

it cannot participate in the disulphide

encoding a modified anaphylactic food a

encoding a natural anaphylactic food al-

Db 901 AGTAGCTAAATCTCCATGCCGTTAAACACACCCGGCCAGTTTGGAGATTTCTTCCCGGC 960

Qy 961 GAGACGCCGAGACAATCATCTTACTTTCAGGGGTTTCAGCAGGAATACGTTGGAGCCGC 1020

Db 961 GAGACGCCGAGACAATCATCTTACTTTCAGGGGTTTCAGCAGGAATACGTTGGAGCCGC 1020

Qy 1021 CTTCAATGCCGAATTCATCAGATACGAGGGTCTGTTAGAGAGATTCAGAGGTTGA 1080

Db 1021 CTTCAATGCCGAATTCATCAGATACGAGGGTCTGTTAGAGAGATTCAGAGGTTGA 1080

Qy 1081 GCAAGAGGAGAGGGCAGAGCGGATGGAGTACTCCGAGTGTGAGAACAAATGAAGGAGT 1140

Db 1081 GCAAGAGGAGAGGGCAGAGCGGATGGAGTACTCCGAGTGTGAGAACAAATGAAGGAGT 1140

Qy 1141 GATAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTTC 1200

Db 1141 GATAGTCAAGTGTCAAGAGGACGCTTGAAGAACTTACTTAAGCAGCGCTAAATCCGCTTC 1200

Qy 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACAAACCCCAATCACTTCAGAGAAGCGGA 1260

Db 1201 AAAGAAAGGCTCCGAAGAGAGGAGATATCACAAACCCCAATCACTTCAGAGAAGCGGA 1260

Qy 1261 GCCGATCTTTCTTAAACAACCTTTGGGAAGTTATTTGAGGTGAACCCAGACAAGAAACCC 1320

Db 1261 GCCGATCTTTCTTAAACAACCTTTGGGAAGTTATTTGAGGTGAACCCAGACAAGAAACCC 1320

Qy 1321 CGAGCTTCAGGACTCGACATGATGTCTACCTGTGTAGAGATCAAAAGAGGAGCTTTGAT 1380

Db 1321 CGAGCTTCAGGACTCGACATGATGTCTACCTGTGTAGAGATCAAAAGAGGAGCTTTGAT 1380

Qy 1381 GCTCCACACCTTCAACTCAAGGCCATGGTTATCTGCTCAACAAAGGAACTCGGAA 1440

Db 1381 GCTCCACACCTTCAACTCAAGGCCATGGTTATCTGCTCAACAAAGGAACTCGGAA 1440

Qy 1441 CCTTGAACCTGCTGCTGTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAAGAGGGA 1500

Db 1441 CCTTGAACCTGCTGCTGTAAGAAAAGAGCAACAACAGAGGGGACGCGGGAAGAGGGA 1500

Qy 1501 GGACGAAGAGCAAGAGAGAGAGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560

Db 1501 GGACGAAGAGCAAGAGAGAGAGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560

Qy 1561 GAAGGAAGCGGATGTCTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620

Db 1561 GAAGGAAGCGGATGTCTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTC 1620

Qy 1621 CGAATCTCATCTGCTTGGCTTCGGTATCAACCGCTGAAACCAACCAAGAAATCTTCTTGC 1680

Db 1621 CGAATCTCATCTGCTTGGCTTCGGTATCAACCGCTGAAACCAACCAAGAAATCTTCTTGC 1680

Qy 1681 AGGTGATAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTC 1740

Db 1681 AGGTGATAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTC 1740

Qy 1741 TGGTCCGGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGAGGATCTCAGCTTGTGAG 1800

Db 1741 TGGTCCGGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGAGGATCTCAGCTTGTGAG 1800

Qy 1801 TGCTCGTCTCAATCTCAATCTCCGTCGTCTCCCTGAGAAAGAGTCTCTCAGAA 1860

Db 1801 TGCTCGTCTCAATCTCAATCTCCGTCGTCTCCCTGAGAAAGAGTCTCTCAGAA 1860

Qy 1861 AGAGGATCAAGAGGAGAAACCAAGAGGGAAGGTTCACTCTTCAATTTGAAGGC 1920

Db 1861 AGAGGATCAAGAGGAGAAACCAAGAGGGAAGGTTCACTCTTCAATTTGAAGGC 1920

Qy 1921 TTTTAAGTGAAGTGAAGCACTCTGATGATTCGATATAGATCAAGCTTTGTACT 1980

Db 1921 TTTTAAGTGAAGTGAAGCACTCTGATGATTCGATATAGATCAAGCTTTGTACT 1980

Qy 1981 CTACTATCAAAAACCTTATCAATAAATAAAGGTTTGTGCGTTGTTCTCC 2032

Db 1981 CTACTATCAAAAACCTTATCAATAAATAAAGGTTTGTGCGTTGTTCTCC 2032

RESULT 5
AAT76613
ID AAT76613 standard; cDNA to mRNA; 2032 BP.
XX
AC AAT76613;
XX
DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hi cDNA clone P41b.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hi; ds.
XX
OS Arachis hypogaea; strain Florunner.
XX
FH Key Location/Qualifiers
FT CDS 50..1930
FT sig_peptide /*tag= a
FT /*tag= 50..115
FT mat_peptide /*tag= b
FT /*tag= 116..1927
FT polyA_signal /*tag= c
FT /*tag= 2005..2010
FT /*tag= d
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US015222.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 04-MAR-1996; 96US-00610424.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
XX
DR WPI: 1997-363453/33.
DR P-PSDB; AAW22150.
XX
PT Peanut allergens Ara hi and Ara hii - used for vaccination and in two-
PT site monoclonal antibody based ELISA.
XX
PS Claim 31; Page 186-188; 354pp; English.
XX
CC This cDNA clone, designated P41b, codes for the major peanut allergen Ara
CC hi (AAW22150), which has multiple IGE binding epitopes (see AAW24165-87).
CC It was amplified from peanut seed cDNA using a primer (see AAT76616)
CC based on an isolated Ara hi peptide (see AAW24206). The sequence shows
CC significant homology with the vicilin family of seed storage proteins of
CC other legumes. The gene is capable of producing a protein product in
CC prokaryotic cells that is recognised by serum IGE from a large proportion
CC of individuals with peanut hypersensitivity. Ara hi and Ara hii (see
CC AAW24164) can be used to raise monoclonal antibodies which are used in a
CC specific two-site MAB ELISA for the detection of Ara hi or Ara hii
CC (claimed). IGE-binding Ara antigen epitopes may be used in vaccines to
CC protect against allergic reactions to peanut allergens, e.g. anaphylactic
CC shock. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 2032 BP; 628 A; 473 C; 530 G; 401 T; 0 U; 0 Other;

Query Match 99.7%; Score 2025.6; DB 2; Length 2032;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AATAATCATATATATTCATCATCATCTATATAGTAGTAGCAGGAGCAATCAGAGGAG 60
Db 1 AATATCATATATATTCATCATCATCTATATAGTAGTAGCAGGAGCAATCAGAGGAG 60

QY	61	GGTTTCTCCAGTATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCA	120
Db			
QY	61	GGTTTCTCCAGTATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTTTCTGCAACGCA	120
Db			
QY	121	TGCCAAGTCAATCACTTTAACAAGAAACACAGAGAAACCCCTGCGCCACAGAGTGCTTCCA	180
Db			
QY	121	TGCCAAGTCAATCACTTTAACAAGAAACACAGAGAAACCCCTGCGCCACAGAGTGCTTCCA	180
Db			
QY	181	GAGTTGCTCAACAGGACCGGATGACTTTGAAGCAAAAGGATCGGATCTCGCTGCACCAA	240
Db			
QY	181	GAGTTGCTCAACAGGACCGGATGACTTTGAAGCAAAAGGATCGGATCTCGCTGCACCAA	240
Db			
QY	241	GCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCA	300
Db			
QY	241	GCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCA	300
Db			
QY	301	AGTTTCCCTCCAGGGGACGGACAGCTGGCGCCCAACCCGAGACTACGATGATGACGG	360
Db			
QY	301	AGTTTCCCTCCAGGGGACGGACAGCTGGCGCCCAACCCGAGACTACGATGATGACGG	360
Db			
QY	361	CGGTCAACCCCGAAGAGAGGAGCGCGATGGGGACCAAGTGGAGCGACCTGGACCGAGGGCGTGA	420
Db			
QY	361	CGGTCAACCCCGAAGAGAGGAGCGCGATGGGGACCAAGTGGAGCGACCTGGACCGAGGGCGTGA	420
Db			
QY	421	AAGAGAAAGACTGGAGACAACCAAGAGAAAGATTGGAGCGACCAAGTCAATCAGCAGCC	480
Db			
QY	421	AAGAGAAAGACTGGAGACAACCAAGAGAAAGATTGGAGCGACCAAGTCAATCAGCAGCC	480
Db			
QY	481	ACGGAATAAGCCCGAAGAGAGAGAGGAAACAAAGAGTGGGAAACACCAAGGTAGCCA	540
Db			
QY	481	ACGGAATAAGCCCGAAGAGAGAGGAAACAAAGAGTGGGAAACACCAAGGTAGCCA	540
Db			
QY	541	TGTGAGGAGAGAAACATCTCGGAAACACCTTCTACTTCCGTCNAAGCGGTTTAGCAC	600
Db			
QY	541	TGTGAGGAGAGAAACATCTCGGAAACACCTTCTACTTCCGTCNAAGCGGTTTAGCAC	600
Db			
QY	601	CCGCTACGGGAACCAAAAACGGTAGATCCGGGTCTCTGACAGAGTTTGACCAAGGTCAAG	660
Db			
QY	601	CCGCTACGGGAACCAAAAACGGTAGATCCGGGTCTCTGACAGAGTTTGACCAAGGTCAAG	660
Db			
QY	661	GCAGTTTCAGAAATCTCCAGAAATCAACCGTATTTGTCAGAGATCGAGGCGCAACCTTAA	720
Db			
QY	661	GCAGTTTCAGAAATCTCCAGAAATCAACCGTATTTGTCAGAGATCGAGGCGCAACCTTAA	720
Db			
QY	721	TGTTCTTCCCAAGCAGCTGATGATTAACATCTTGTATCCAGAGGCGCAAGCCAC	780
Db			
QY	721	TGTTCTTCCCAAGCAGCTGATGATTAACATCTTGTATCCAGAGGCGCAAGCCAC	780
Db			
QY	781	CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACT	840
Db			
QY	781	CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACT	840
Db			
QY	841	CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAAGAACTTCAG	900
Db			
QY	841	CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAAGAACTTCAG	900
Db			
QY	901	AGTAGTAAATCTCCATGCGGTTTAAACACACCCCGCAGTTTGAGGATTTCTTCCCGGC	960
Db			
QY	901	AGTAGTAAATCTCCATGCGGTTTAAACACACCCCGCAGTTTGAGGATTTCTTCCCGGC	960
Db			
QY	961	GAGCAGCGGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACGTTGGAGGCGGC	1020
Db			
QY	961	GAGCAGCGGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACGTTGGAGGCGGC	1020
Db			
QY	1021	CTTCAATCGGAAATTCATAGATACGAGGAGTGCTGTTAGAAAGAAATGCAAGGAGTGA	1080
Db			
QY	1021	CTTCAATCGGAAATTCATAGATACGAGGAGTGCTGTTAGAAAGAAATGCAAGGAGTGA	1080
Db			
QY	1081	GCAAGAGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTACTGAGAACAAATGAAGGAGT	1140
Db			
QY	1081	GCAAGAGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTACTGAGAACAAATGAAGGAGT	1140
Db			

QY	1141	GATAGTCAAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTC	1200
Db			
QY	1141	GATAGTCAAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCGTTCTC	1200
Db			
QY	1201	AAAGAAAGGCTCCGAAAGAGAGGAGATATCACCACCCCAATCAACTTTGAGAGAAGGGA	1260
Db			
QY	1201	AAAGAAAGGCTCCGAAAGAGAGGAGATATCACCACCCCAATCAACTTTGAGAGAAGGGA	1260
Db			
QY	1261	GCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCC	1320
Db			
QY	1261	GCCGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCC	1320
Db			
QY	1321	CCAGTTCCAGGACCTGGACATGCTCACCTCTGTGTAGAGATCAAAGAGAGCTTTTGAT	1380
Db			
QY	1321	CCAGTTCCAGGACCTGGACATGCTCACCTCTGTGTAGAGATCAAAGAGAGCTTTTGAT	1380
Db			
QY	1381	GCTCCACACTTTCAACTCAAAGGCCATGGTTATCGTCTGTCACAAAGGAACCTGGAAA	1440
Db			
QY	1381	GCTCCACACTTTCAACTCAAAGGCCATGGTTATCGTCTGTCACAAAGGAACCTGGAAA	1440
Db			
QY	1441	CCTTGAACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGGAAGAGAGA	1500
Db			
QY	1441	CCTTGAACTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGGAAGAGAGA	1500
Db			
QY	1501	GGACGAAGCAAGAGAGGAGGGAAGTAACAAGAGAGTGGTGTAGGTACACAGCGAGGTT	1560
Db			
QY	1501	GGACGAAGCAAGAGAGGAGGGAAGTAACAAGAGAGTGGTGTAGGTACACAGCGAGGTT	1560
Db			
QY	1561	GAAAGAGCGGATGTTTCATCATGCGCAGAGCTCATCAGTAGCCATCAACGCTTCTCTC	1620
Db			
QY	1561	GAAAGAGCGGATGTTTCATCATGCGCAGAGCTCATCAGTAGCCATCAACGCTTCTCTC	1620
Db			
QY	1621	CGAATCCATCTGCTTGGCTTCGATCAACCGCTGAAACCAACACAGAACTTCTCTTGC	1680
Db			
QY	1621	CGAATCCATCTGCTTGGCTTCGATCAACCGCTGAAACCAACACAGAACTTCTCTTGC	1680
Db			
QY	1681	AGGTGATTAAGGACAATGTGTAGACCCAGATAGAGAAGCAAGCGAAGGATTTAGCATTC	1740
Db			
QY	1681	AGGTGATTAAGGACAATGTGTAGACCCAGATAGAGAAGCAAGCGAAGGATTTAGCATTC	1740
Db			
QY	1741	TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCCAGAGGAATCTCACTTTGTGAG	1800
Db			
QY	1741	TGGTTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCCAGAGGAATCTCACTTTGTGAG	1800
Db			
QY	1801	TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCGAAGAGAGTCTCTCTGAGAA	1860
Db			
QY	1801	TGCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCGAAGAGAGTCTCTCTGAGAA	1860
Db			
QY	1861	AGAGATCAAGAGGAGGAAAAACCAAGGAGGAAAGGTCCTCTTTCAATTTTGAAGGC	1920
Db			
QY	1861	AGAGATCAAGAGGAGGAAAAACCAAGGAGGAAAGGTCCTCTTTCAATTTTGAAGGC	1920
Db			
QY	1921	TTTTAAGTGAAGTGGAGGCACTTGTATGTATCGATAATAAGATCAAGCTTTGTACT	1980
Db			
QY	1921	TTTTAAGTGAAGTGGAGGCACTTGTATGTATCGATAATAAGATCAAGCTTTGTACT	1980
Db			
QY	1981	CTACTATCAAAAACTTATCAATAATAAAAAACGTTTGTGCTGTTTCTCC	2032
Db			
QY	1981	CTACTATCAAAAACTTATCAATAATAAAAAACGTTTGTGCTGTTTCTCC	2032
Db			

RESULT 6

AAZ06382
ID AAZ06382 standard; DNA; 1930 BP.

XX AAZ06382;

XX 17-OCT-2003 (revised)

XX 09-NOV-1999 (first entry)

XX Peanut allergen, Ara h 1.

KW allergy; immune response; transgenic; allergen; epitope;

immunoglobulin E; Ig E; binding site; peanut; ds.

Arachis hypogaea.

Key Location/Qualifiers
CDS 50..1930
/*tag= a
/product= "Ara h 1"

WO9938978-A1.

05-AUG-1999.

29-JAN-1999; 99WO-US002031.

31-JAN-1998; 98US-0073283P.

13-FEB-1998; 98US-0074590P.

13-FEB-1998; 98US-0074624P.

13-FEB-1998; 98US-0074633P.

27-AUG-1998; 98US-00141220.

XX (UYAR-) UNIV ARKANSAS.

PA (UNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

PA (SOSI/) SOSIN H.

XX Sosin H, Bannon GA, Burks AW, Sampson HA;

XX WPI; 1999-479189/40.

DR P-PSDB; AAY15244.

XX Modified allergen with reduced IgE binding, useful for treating e.g.

PT allergies.

XX Disclosure; Page 32-33; 46pp; English.

XX This is the nucleotide sequence of the Ara h 1 protein from Arachis
CC hypogaea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding
CC epitopes, four of which are immunodominant (AAY15247, AAY15249, AAY15250
CC and AAY15263). By modifying the IgE binding sites the ability of the
CC allergen to provoke an immune response is downregulated. The epitopes of
CC the IgE binding sites can therefore be modified in genetically engineered
CC plants and animals to elicit less of an allergic response. (Updated on 17
CC -OCT-2003 to standardise OS field)

XX Sequence 1930 BP; 595 A; 455 C; 514 G; 366 T; 0 U; 0 Other;

Query Match 95.0%; Score 1930; DB 2; Length 1930;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1930; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGGGAG 60

DB 1 AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGGGAG 60

QY 61 GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTGGCTTCAGTTCTCAGACGCA 120

DB 61 GGTTTCTCCACTGATGCTGTGTAGGGATCCTTGCTGGCTTCAGTTCTCAGACGCA 120

QY 121 TGCCAAGTCTATCACTTACCAGAAAGAAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA 180

DB 121 TGCCAAGTCTATCACTTACCAGAAAGAAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA 180

QY 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCGATCGAGTCTCGCTGCCACCA 240

DB 181 GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCGATCGAGTCTCGCTGCCACCA 240

QY 241 GCTCGAGTATGATCCTCGTTGTCTATGATCTTCGAGGACACATGGGACCAACCAACCA 300

DB 241 GCTCGAGTATGATCCTCGTTGTCTATGATCTTCGAGGACACATGGGACCAACCAACCA 300

QY 301 ACCTTCCCTCCAGGGAGGGGACACGTGGCGCCCAACCCGGAGACTTACGATGATGACCG 360

DB 301 ACCTTCCCTCCAGGGAGGGGACACGTGGCGCCCAACCCGGAGACTTACGATGATGACCG 360

QY 361 CCGTCAACCCCGAAGAGAGAGAGAGCCGATGGGACCAGCTGGACCCAGGGAGCGTGA 420
DB 361 CCGTCAACCCCGAAGAGAGAGAGAGCCGATGGGACCAGCTGGACCCAGGGAGCGTGA 420
QY 421 AAGAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGTCAATCAGCAGCC 480
DB 421 AAGAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGTCAATCAGCAGCC 480
QY 481 ACGGAAATTAAGCCCGAAGAGAGAGAGAGAAAGAGATGGGGAAACCAACAGGTAGCCA 540
DB 481 ACGGAAATTAAGCCCGAAGAGAGAGAGAGAAAGAGATGGGGAAACCAACAGGTAGCCA 540
QY 541 TGTGAGGGAAGAAAATCTCGGAAACAAACCTTTCTACTTCCCGTCAAGGGCGTTTAGCAC 600
DB 541 TGTGAGGGAAGAAAATCTCGGAAACAAACCTTTCTACTTCCCGTCAAGGGCGTTTAGCAC 600
QY 601 CCGCTACGGGAACCAAAACGGTAGGATCGGGTCTGTGACAGAGTTTGACCAAGGTCAAG 660
DB 601 CCGCTACGGGAACCAAAACGGTAGGATCGGGTCTGTGACAGAGTTTGACCAAGGTCAAG 660
QY 661 GCAGTTTCAGAAATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT 720
DB 661 GCAGTTTCAGAAATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCT 720
QY 721 TGTTCCTCCAGACCGCTGATGATTAACATCTTGTATTCAGCAAGGGCAAGCCAC 780
DB 721 TGTTCCTCCAGACCGCTGATGATTAACATCTTGTATTCAGCAAGGGCAAGCCAC 780
QY 781 CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACAGGGCCATCACT 840
DB 781 CGTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACAGGGCCATCACT 840
QY 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAACAGAACTCAG 900
DB 841 CAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACCAACAGAACTCAG 900
QY 901 AGTAGTAAATCTCCATGCCGTTTAAACACCCGGCCAGTTTGAGGATTTCTTCCGGC 960
DB 901 AGTAGTAAATCTCCATGCCGTTTAAACACCCGGCCAGTTTGAGGATTTCTTCCGGC 960
QY 961 GAGCAGCCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGCCG 1020
DB 961 GAGCAGCCGAGACCAATCATCTTCTGAGGGCTTTCAGCAGGAATACGTTGGAGGCCG 1020
QY 1021 CTTCAATCGGAATTCATGAGATACGGAGGTGCTGTTAGAGAGAAATGCAGGAGTGA 1080
DB 1021 CTTCAATCGGAATTCATGAGATACGGAGGTGCTGTTAGAGAGAAATGCAGGAGTGA 1080
QY 1081 GCAAGAGGAGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT 1140
DB 1081 GCAAGAGGAGAGAGGGCAGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGT 1140
QY 1141 GATAGTCAAGGTGTCAGAGGACAGCTTGAAGAACTTACTTAAGCAGCTAAATCCGCTTC 1200
DB 1141 GATAGTCAAGGTGTCAGAGGACAGCTTGAAGAACTTACTTAAGCAGCTAAATCCGCTTC 1200
QY 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATCAACCAACCAATCACTTGAGAGAGAGCCGA 1260
DB 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATCAACCAACCAATCACTTGAGAGAGAGCCGA 1260
QY 1261 GCCGATCTTTCTAAACACTTTGGGAAGTTATTGAGGTGAAGCCAGACAAAGAAACCC 1320
DB 1261 GCCGATCTTTCTAAACACTTTGGGAAGTTATTGAGGTGAAGCCAGACAAAGAAACCC 1320
QY 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGGAGCTTTGAT 1380
DB 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGGAGCTTTGAT 1380
QY 1381 GCTCCCACTTCACTCAAAGGCCATGTTATCGTCGTCGTCAACAAAGGAACTGGA 1440
DB 1381 GCTCCCACTTCACTCAAAGGCCATGTTATCGTCGTCGTCAACAAAGGAACTGGA 1440

QY	1441	CCTTGAACCTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGA	1500
Db	1441	CCTTGAACCTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGA	1500
QY	1501	GGACGAAGACGAAGAGAGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
Db	1501	GGACGAAGACGAAGAGAGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT	1560
QY	1561	GAAGGAAGCGGATGTCTCATCATGCGAGCGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
Db	1561	GAAGGAAGCGGATGTCTCATCATGCGAGCGCTCATCCAGTAGCCATCAACGCTTCCTC	1620
QY	1621	CGAATCCATCTCTTGGCTTCGGTATCAACGCTGAAAAACCAACAGAGATCTTCCTGC	1680
Db	1621	CGAATCCATCTCTTGGCTTCGGTATCAACGCTGAAAAACCAACAGAGATCTTCCTGC	1680
QY	1681	AGGTGATAGGCAATGTGATAGACAGATAGAGAAAGCAAGCGAGGATTTAGCATTCCT	1740
Db	1681	AGGTGATAGGCAATGTGATAGACAGATAGAGAAAGCAAGCGAGGATTTAGCATTCCT	1740
QY	1741	TGGTCCGGTGAACAAAGTTGAGAAGCTCATCAAAAACCAAGGAATCTCACTTTGTGAG	1800
Db	1741	TGGTCCGGTGAACAAAGTTGAGAAGCTCATCAAAAACCAAGGAATCTCACTTTGTGAG	1800
QY	1801	TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCCTGAGAA	1860
Db	1801	TGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCCTGAGAA	1860
QY	1861	AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCACTCCTTCAATTTGAGGC	1920
Db	1861	AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGTCCACTCCTTCAATTTGAGGC	1920
QY	1921	TTTTAACTGA 1930	
Db	1921	TTTTAACTGA 1930	
RESULT 7			
AAAT76612			
ID	AAAT76612 standard; cDNA to mRNA, 1949 BP.		
XX	AAAT76612;		
AC	AAAT76612;		
XX	17-OCT-2003 (revised)		
DT	29-DEC-1997 (first entry)		
XX	Peanut allergen Ara hI cDNA clone P17.		
DE	Peanut; seed storage protein; allergen; allergy; hypersensitivity;		
XX	vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;		
KW	ELISA; analysis; Ara hI; ds.		
KW	Arachis hypogaea; strain Florunner.		
OS	Key		
XX	Location/Qualifiers		
XX	3. .1847		
FT	CDS		
FT	/*tag= a		
FT	3. .68		
FT	sig_peptide		
FT	/*tag= b		
FT	65. .1844		
FT	mat_peptide		
FT	/*tag= c		
FT	1918. .1923		
FT	polya_signal		
FT	/*tag= d		
XX	WO9724139-A1.		
PN	10-JUL-1997.		
XX	23-SEP-1996;		
XX	96WO-US015222.		
XX	29-DEC-1995;		
PR	95US-0009455P.		
PR	04-MAR-1996;		
XX	96US-0061042A.		

PA	(UYAR-) UNIV ARKANSAS.	
XX	Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;	
PI	WPI; 1997-363453/33.	
XX	P-PSDB; AAW22149.	
DR	Peanut allergens Ara hI and Ara hII - used for vaccination and in two-	
DR	site monoclonal antibody based ELISA.	
XX	Claim 31; Page 183-185; 354pp; English.	
PT	This cDNA clone, designated p17, codes for the major peanut allergen Ara	
XX	hI (AAW22149), which has multiple IGE binding epitopes (see AAW24165-87).	
CC	It was amplified from peanut seed cDNA using a primer (see AAT76616)	
CC	based on an isolated Ara hI peptide (see AAW24206). The sequence shows	
CC	significant homology with the vicilin family of seed storage proteins of	
CC	other legumes. The gene is capable of producing a protein product in	
CC	prokaryotic cells that is recognised by serum IGE from a large proportion	
CC	of individuals with peanut hypersensitivity. Ara hI and Ara hII (see	
CC	AAW24164) can be used to raise monoclonal antibodies which are used in a	
CC	specific two-site Mab ELISA for the detection of Ara hI or Ara hII	
CC	(claimed). IGE-binding Ara antigen epitopes may be used in vaccines to	
CC	protect against allergic reactions to peanut allergens, e.g. anaphylactic	
CC	shock. (Updated on 17-OCT-2003 to standardise OS field)	
XX	Sequence 1949 BP; 599 A; 455 C; 517 G; 378 T; 0 U; 0 Other;	
QY	Query Match 86.2%; Score 1752.4; DB 2; Length 1949;	
Db	Best Local Similarity 95.0%; Pred. No. 0;	
QY	Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;	
QY	48 CAATGAGAGGAGGGTTCCTCCACTGATGCTGTGCTAGGATCTCTGCTGGCTTCAG 107	
Db	1 CAATGAGAGGAGGGTTCCTCCACTGATGCTGTGCTAGGATCTCTGCTGGCTTCAG 60	
QY	108 TTTCTGCAACGATGCCAAGTCATCACTTACCAAGAAAGAAACAGAGAACCCCTCGGCC 167	
Db	61 TTTCTGCAACGAGGCCCAAGTCACC-----TTACCGGAAAGAAACAGAGAACCCCTCGGCC 114	
QY	168 AGAGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGT 227	
Db	115 AGAGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGGAGT 174	
QY	228 CTCGCTGCACCAAGCTCGAGTATGATCTCTGCTGTGCTATGATCTCTCGAGGACACATG 287	
Db	175 CTCGCTGCACCAAGCTCGAGTATGATCTCTGCTGTGCTATG-----ACACTG 222	
QY	288 GCACCAACCAACCAAGTTCCCTCCAGGGAGCGGACACGTGGCGCCCAACCCGAGACT 347	
Db	223 GCGCCACCAACCAACGTCACCTCCAGGGAGCGGACACGTGGCGCCCAACCCGAGACT 282	
QY	348 AGGATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGACCCAGCTGGAC 407	
Db	283 AGGATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGGATGGGACCCAGCTGGAC 342	
QY	408 CGAGGAGCGTGAAGAGAGAGAAAGACTGGAGACCAACCAAGAGAGATTGGAGGCCACCA 467	
Db	343 CGAGGAGCGTGAAGAGAGAGAAAGACTGGAGACCAACCAAGAGAGATTGGAGGCCACCA 402	
QY	468 GTTCATCAGACGCCACGGAAAATAAGGCCCGAAGGAAGAGAGAGAAACAAGATGGGAA 527	
Db	403 GTTCATCAGACGCCACGGAAAATAAGGCCCGAAGGAAGAGAGAGAAACAAGATGGGAA 462	
QY	528 CACAGGTAGCCATGTGAGGGAAGAAACATCTCCGAACCAACCTTTCTTCTTCCGTCAA 587	
Db	463 CACAGGTAGCCATGTGAGGGAAGAAACATCTCCGAACCAACCTTTCTTCTTCCGTCAA 522	
QY	588 GCGGTTTTAGCACCCGCTACCGGAACCAAAACCGGTAGGATCCGGTCTCTGCAGAGGTTG 647	
Db	523 GCGGTTTTAGCACCCGCTACCGGAACCAAAACCGGTAGGATCCGGTCTCTGCAGAGGTTG 582	
QY	648 ACCAAAGGTCAAGGCAGGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCGAGATCGAGGCCA 707	

||||| 583 ACCAAGGTCAAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTTGTGCAGATCGAGCCCA 642
Qy ||||| 708 AACCTAAACATCTTTGTTCTTCCCAAGCAGCGTGATGCTGATAAATCATCTTTGTTATCCAGC 767
Db ||||| 643 GACCTAAACATCTTTGTTCTTCCCAAGCAGCGTGATGCTGATAAATCATCTTTGTTATCCAGC 702
Qy ||||| 768 AAGGGCAAGCCACCGTCACCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAAATCTTTGAGC 827
Db ||||| 703 AAGCAAGCCACCGTGACCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAAATCTTTGAGC 762
Qy ||||| 828 AGGGCCATGCACTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATGACA 887
Db ||||| 763 AGGGCCATGCACTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAATCGACATGACA 822
Qy ||||| 888 ACAGAAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACACCGCGCAGTTTGGAG 947
Db ||||| 823 ACCAGAACTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACCGCGCGCAGTTTGGAG 882
Qy ||||| 948 ATTCTTCCCGCGGAGCAGCGGAGACCAATCATCTTCTTACCTTGCAGGGCTTCAGCAGGAATA 1007
Db ||||| 883 ATTCTTCCCGCGGAGCAGCGGAGACCAATCATCTTCTTACCTTGCAGGGGATTGACAGGGAATA 942
Qy ||||| 1008 CGTTGAGGCGCGCTTCAATGCGGAATTCATGAGATACGAGGGTCTGTAGAAGAGA 1067
Db ||||| 943 CTTTGGAGCGCGCTTCAATGCGGAATTCATGAGATACGAGGGTCTGTAGAAGAGA 1002
Qy ||||| 1068 ATGAGGAGGTGAGCAAGAGAGAGAGGGGAGAGGGGATGAGTACTCGGAGTAGTGAGA 1127
Db ||||| 1003 ATGAGGAGGAGAGCAAGAGAGAGAGGGGAGAGGGGATGAGTACTCGGAGTAGTG- 1059
Qy ||||| 1128 ACATGAGGAGGTGATGATCAAGTGTCAAGAGGACGCTTGAAGAACTTACTTAAGCAG 1187
Db ||||| 1060 ATAATGAAGAGGTGATGATCAAGTGTCAAGAGGACGCTTCAAGAACTTACTTAAGCAG 1119
Qy ||||| 1188 CTAATCCGTCTCAAGAAAGGCTCCGAGAGAGGAGATATCACCAACCAATCAACT 1247
Db ||||| 1120 CTAATCCGTCTCAAGAAAGGCTCC- - - GAAGAGGAGATATCACCAACCAATCAACT 1176
Qy ||||| 1248 TGAGAGAGCGGAGCGCGATCTTTCTTAACTTTGGGAAGTTATTGAGGTGAAGCCAG 1307
Db ||||| 1177 TGAGAGAGTGGGAGCGCGATCTTTCTTAACTTTGGGAGTTATTGAGGTGAAGCCAG 1236
Qy ||||| 1308 ACAAGAGAACCCCGAGCTTCAGACCTGACATGATGCTCAGCTGTGTGATAGATCAAG 1367
Db ||||| 1237 ACAAGAGAACCCCGAGCTTCAGACCTGACATGATGCTCAGCTGTGTGATAGATCAAG 1296
Qy ||||| 1368 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATGCTCGTCAACA 1427
Db ||||| 1297 AAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATGCTCGTCAACA 1356
Qy ||||| 1428 AAGGAACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGCG 1487
Db ||||| 1357 AAGGAACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGCG 1416
Qy ||||| 1488 GGGAA- - - - - GAAGAGGAGGACGAGACGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1538
Db ||||| 1417 GGGAAACAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1476
Qy ||||| 1539 TGGTGTAGGTACAGCGAGGTTGAAGGAGGCGATGTTTCACTGTCAGCAGAGCTCATC 1598
Db ||||| 1477 TGGTGTAGGTACAGCGAGGTTGAAGGAGGCGATGTTTCACTGTCAGCAGAGCTCATC 1536
Qy ||||| 1599 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1658
Db ||||| 1537 CAGTAGCCATCAACGCTTCCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAA 1596
Qy ||||| 1659 ACAACACAGAAATCTTCTTGCAGGTGATGAAGCAATGTATGATAGACAGATAGAAGC 1718
Db ||||| 1597 ACAACACAGAAATCTTCTTGCAGGTGATGAAGCAATGTATGATAGACAGATAGAAGC 1656
Qy ||||| 1719 AAGCGAAGGATTTAGCAATTCCTCGGTGGGTGCAAGTTGAGAGGCTCATCAAAACC 1778

Db 1657 AAGCGAAGGATTTAGCATTCCTGGTTCGGTGCAACAAGTTGAGAAGCTCATCAAAACC 1716
Qy 1779 AGAAGAAATCTCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCAATCTCCTGCTC 1838
Db 1717 AGAGGGAGTCTCACTTTGTGAGTGTGCTCCTCAATCTCAATCTCCTGCTC- - - - - 1765
Qy 1839 CTGAGAAAGAGTCTCCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTC 1898
Db 1766 - - - - - GTCTCTGAAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTC 1815
Qy 1899 CACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAATCGAGGCAACTTGTATTATCTATCGAT 1958
Db 1816 CACTCCTTTCAATTTTGAAGGCTTTTAACTGAGAATCGAGGAACTTGTATTATCTATCAT 1875
Qy 1959 AATAAGATCAGCGTTTGTACTACTACTATCCAAAAAATTATCAATAAATAAAACGTTTG 2018
Db 1876 AATAAGATCAGCGTTTGTAAATCTACTATCCAAAAAATTATCAATAAATAAAACGTTTG 1935
Qy 2019 TCGGTTGTTTCTCC 2032
Db 1936 TCGGTTGTTTCTCC 1949

RESULT 8
ABX70604
ID ABX70604 standard; cDNA; 1949 BP.
XX
AC ABX70604;
XX AC
DT 26-MAR-2003 (first entry)
XX
DE Peanut Ara h1 cDNA clone P17.
XX
KW Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; IgE binding site;
KW anaphylactic food allergen; anti-allergenic; vaccine; wound healing.
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabin PA, Shin DS, Stanley JS;
XX
DR WPI: 2003-018765/01.
DR P-PSDB; ABUS2413.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Example 3; Fig 11; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified

CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a cDNA
CC encoding a peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 1949 BP; 599 A; 455 C; 517 G; 378 T; 0 U; 0 Other;

Query Match 86.2%; Score 1752.4; DB 8; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;

QY	48	CAATGAGAGGGGGTTTCTCCACTGATGCTGTGTGCTAGGGATCCTTTGCTCGGCTTCAG	107
DB	1	CAATGAGAGGGGGTTTCTCCACTGATGCTGTGTGCTAGGGATCCTTTGCTCGGCTTCAG	60
QY	108	TTTCTGCAACGATGCAAGTCAATCACTTACAGAGAAACACAGAGAACCCCTGCGCC	167
DB	61	TTTCTGCAACGAGGGCAAGTCAAC-----TTACCGGAAACACAGAGAACCCCTGCGCC	114
QY	168	AGAGGTGCTCCAGAGTGTCTAACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGT	227
DB	115	AGAGGTGCTCCAGAGTGTCTAACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGT	174
QY	228	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGACACACTG	287
DB	175	CTCGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATG-----ACACTG	222
QY	288	GCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCGCCCAACCCGAGACT	347
DB	223	GGCCCAACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCGCCCAACCCGAGACT	282
QY	348	ACGATGATGACCGCGCTCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	407
DB	283	ACGATGATGACCGCGCTCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	342
QY	408	CGAGGAGCGCTGAAG	467
DB	343	CGAGGAGCGCTGAAG	402
QY	468	GTCAATCAGAGGCCACGGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	527
DB	403	GTCAATCAGAGGCCACGGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	462
QY	528	CACGAGTAGCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	587
DB	463	CACGAGTAGCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	522
QY	588	GGCGGTTTAGCACCCGCTACGGGAGCAAAACGGTAGGATCCGGGTCTCGCAGAGGTTG	647
DB	523	GGCGGTTTAGCACCCGCTACGGGAGCAAAACGGTAGGATCCGGGTCTCGCAGAGGTTG	582
QY	648	ACCAAGGTCAAGCGAGTTTCAGATCTCCAGATCACCGTATGTGCGAGATCGAGGCCA	707
DB	583	ACCAAGGTCAAGCGAGTTTCAGATCTCCAGATCACCGTATGTGCGAGATCGAGGCCA	642
QY	708	AACTTAACACTCTGTCTTCCAGCAGCGCTGATGAATCAATCTCTGTTTATCCAGC	767
DB	643	AACTTAACACTCTGTCTTCCAGCAGCGCTGATGAATCAATCTCTGTTTATCCAGC	702
QY	768	AAGGGCAAGCCACCGTGAACGGTGAAGAAATGGCAATTAACAGAAAGAGCTTTAATCTTGACG	827
DB	703	AAGGGCAAGCCACCGTGAACGGTGAAGAAATGGCAATTAACAGAAAGAGCTTTAATCTTGACG	762
QY	828	AGGGCATTGCACTCAGAAATCCATCGGTTTCATTTCTTACATCTGTGAACCCGCCATGACA	887
DB	763	AGGGCATTGCACTCAGAAATCCATCGGTTTCATTTCTTACATCTGTGAATCGACATGACA	822
QY	888	ACCAGAACCTCAGAGTAGCTAAATCTCCATCCCGTTTAAACACACCCGCCAGTTTGGG	947
DB	823	ACCAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCCAGTTTGGG	882

QY	948	ATTTCTTCCCGCGCAGCGAGCCGAGACCAATCATCTACTTGCAGGGCTTTCAGCAGGAATA	1007
DB	883	ATTTCTTCCCGCGCAGCGAGCCGAGACCAATCATCTACTTGCAGGGATTCAGCAGGAATA	942
QY	1008	CGTTGGAGCGCCCTTCAATGCGGAATTCATGAGATAGCGAGGGTGTCTTTAGAGAGA	1067
DB	943	CGTTGGAGCGCCCTTCAATGCGGAATTCATGAGATAGCGAGGGTGTCTTTAGAGAGA	1002
QY	1068	ATGAGGAGGAGTTCAGCAAGAGGAGAGCGGAGCGGATGGAGTACTCGGAGTAGTGAGA	1127
DB	1003	ATGAGGAGGAGTTCAGCAAGAGGAGAGCGGAGCGGATGGAGTACTCGGAGTAGTG---	1059
QY	1128	ACAATGAAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCAG	1187
DB	1060	ATAATGAAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCAG	1119
QY	1188	CTAAATCCGTCTCAAAGAAAGGCTCCGAAAGAGAGGGAGATATCACCAACCAATCAACT	1247
DB	1120	CTAAATCCGTCTCAAAGAAAGGCTCC---GAAGAGGAAGATATCACCAACCAATCAACT	1176
QY	1248	TGAGAGAAAGGCGAGCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAG	1307
DB	1177	TGAGAGATGGCGAGCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAG	1236
QY	1308	ACAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAG	1367
DB	1237	ACAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAG	1296
QY	1368	AAGGAGCTTTGATGTCTCCACACTTTCAAACCTCAAAGGCCATGGTTATCGTCTGTCAACA	1427
DB	1297	AAGGAGCTTTGATGTCTCCACACTTTCAAACCTCAAAGGCCATGGTTATCGTCTGTCAACA	1356
QY	1428	AAGGAACCTGGAACCTTGAACCTCGTGGCTGTGAAGAAAAGAGCAACAACAGAGGGGACCG	1487
DB	1357	AAGGAACCTGGAACCTTGAACCTCGTGGCTGTGAAGAAAAGAGCAACAACAGAGGGGACCG	1416
QY	1488	GGGA-----GAAGAGGAGGACGAAGACGAAGAGAGAGAGAGAGAGAGAGAGAG	1538
DB	1417	GGGAACAAGAGTGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1476
QY	1539	TGCGTAGGTACACAGCAGGTTGAAGAAAGGCGATGTGTTTCATCATGCGCAGCGCTCATC	1598
DB	1477	TGCGTAGGTACACAGCAGGTTGAAGAAAGGCGATGTGTTTCATCATGCGCAGCGCTCATC	1536
QY	1599	CAGTAGGCATCAACGCTTCTCCGAACCTCATCTGTTGGCTTCGGTATCAACGCTGAAA	1658
DB	1537	CAGTAGGCATCAACGCTTCTCCGAACCTCATCTGTTGGCTTCGGTATCAACGCTGAAA	1596
QY	1659	ACAACACAGAAATCTTCTTGCGGTGATAAGGACAATGTGATAGACCAGATAGAGAAGC	1718
DB	1597	ACAACACAGAAATCTTCTTGCGGTGATAAGGACAATGTGATAGACCAGATAGAGAAGC	1656
QY	1719	AAGCGAAGGATTTAGCATTCCTCGGTGCGGTGAACAAGTTGAGNAGCTCATCAAAAACC	1778
DB	1657	AAGCGAAGGATTTAGCATTCCTCGGTGCGGTGAACAAGTTGAGNAGCTCATCAAAAACC	1716
QY	1779	AGAAGGAATCTCACCTTCTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGTCCTC	1838
DB	1717	AGAAGGAATCTCACCTTCTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGTC-----	1765
QY	1839	CTGAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGAGGTC	1898
DB	1766	-----GTCTCTGAAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGAGGTC	1815
QY	1899	CATCTCTTTCAATTTTGAAGCTTTTAACTGAGAAATGGAGGCACTTGTATGTATCGAT	1958
DB	1816	CATCTCTTTCAATTTTGAAGCTTTTAACTGAGAAATGGAGGCACTTGTATGTATCGAT	1875
QY	1959	AATAAGATCAGCTTTTGTCTACTATCCAAAACCTTATCAATAAATAAAGAGGTTTG	2018
DB	1876	AATAAGATCAGCTTTTGTATCTACTATCCAAAACCTTATCAATAAATAAAGAGGTTTG	1935
QY	2019	TGCGTTGTTTCTCC	2032

Db 523 |||||GGCGGTTTAGCACCGCTAGCGGAAACCAAAACGGTAGGATCCGCGCTCTCCAGAGGTTTG 582
Qy 648 ACCAAAGGTCAGGCGAGTTTTCAGAACTCCAGAACTCACCGTATTGTGCGAGATCGAGGCCA 707
Db 583 ACCAAAGGTCAAGACGAGTTTTCAGAACTCCAGAACTCACCGTATTGTGCGAGATCGAGGCCA 642
Qy 708 AACCTAACACTCTTGTCTTCTCCCAAGCAGCGTGATGCTGATAACATCTCTGTTTATCCAGC 767
Db 643 GACCTAACACTCTTGTCTTCTCCCAAGCAGCGTGATGCTGATAACATCTCTGTTTATCCAGC 702
Qy 768 AAGGCGAAGCCACCGTAGCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGAGC 827
Db 703 AAGGCAAGCCACCGTAGCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGAGC 762
Qy 828 AGGGCCATGCACCTCAGAACTCCCATCCGTTTCAATTTCTTACATCTTGAACCCGCAATGACA 887
Db 763 AGGGCCATGCACCTCAGAACTCCCATCCGTTTCAATTTCTTACATCTTGAATCCACATGACA 822
Qy 888 ACCAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCGCCACGTTTGAGG 947
Db 823 ACCAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACGCGCCGCCAGTTTGAGG 882
Qy 948 ATTTCTTCCGGGAGCAGCGAGCAATCATCTTACATCTTGAGGCTTCAGAGGAATA 1007
Db 883 ATTTCTTCCGGGAGCAGCGAGCAATCATCTTACATCTTGAGGGAATTCAGAGGAATA 942
Qy 1008 CGTTGAGCGCGCTTCAATTCGGGAATTCATAGATACGGAGGTCGTGTTAGAGAGA 1067
Db 943 CTTTGAGCGCGCTTCAATTCGGGAATTCATAGATACGGAGGTCGTGTTAGAGAGA 1002
Qy 1068 ATGCGAGGTTGAGCAAGAGGAGAGAGGCGCAGAGCGATGCGAGTAGCTCGGAGTAGTGAGA 1127
Db 1003 ATGCGAGGAGAGCAAGAGGAGAGAGGCGCAGAGCGATGCGAGTAGCTCGGAGTAGTG--- 1059
Qy 1128 ACAATGAAGAGTAGTAGTCAAAGTGTCAAAGAGCACTGTAAGAGCACTTACTAAGCAGC 1187
Db 1060 ATAATGAAGAGTAGTAGTCAAAGTGTCAAAGAGCAGCGTTCAAGAACTTACTAAGCAGC 1119
Qy 1188 CTAAATCCGTCTCAAAGAAAGGCTCGAAGAGAGAGGAGATATCAACCAACCAATCAACT 1247
Db 1120 CTAAATCCGTCTCAAAGAAAGGCTCC---GAAGAGGAAGATATCAACCAACCAATCAACT 1176
Qy 1248 TGAGAGAGGCGAGCGCCGATCTTTCTAACAACTTTGGGAAGTTATTGAGGTGAAGCCAG 1307
Db 1177 TGAGAGATGCGAGCGCCGATCTTTCTAACAACTTTGGGAGGTTATTGAGGTGAAGCCAG 1236
Qy 1308 ACAAGAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG 1367
Db 1237 ACAAGAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAG 1296
Qy 1368 AAGGAGCTTTGATGCTCCACACTTCACTCAAAGGCCATGTTATCGTCTGTCACCA 1427
Db 1297 AAGGAGCTTTGATGCTCCACACTTCACTCAAAGGCCATGTTATCGTCTGTCACCA 1356
Qy 1428 AAGGAACTGGAAACCTTGAACCTCGTAGCTGTGAAGAAAGAGCAACACAGAGGGGACGCG 1487
Db 1357 AAGGAACTGGAAACCTTGAACCTCGTAGCTGTGAAGAAAGAGCAACACAGAGGGGACGCG 1416
Qy 1488 GGGAA-----GGAAGAGGAGCGAAGACGAGAGAGGAGGAAGTAACAGAGG 1538
Db 1417 GGGAAACAGAGTGGGAAGAGAGGAGGAAGATGAAGAGAGGAGGGAAGTAACAGAGG 1476
Qy 1539 TCGGTAGGTACACAGCGAGGTTGAAGAGGCGATGTTTCATCATGCCAGAGCTCATC 1598
Db 1477 TCGGTAGGTACACAGCGAGGTTGAAGAGGCGATGTTTCATCATGCCAGAGCTCATC 1536
Qy 1599 CAGTAGCCATCAACCGTTCTCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAA 1658
Db 1537 CAGTAGCCATCAACCGTTCTCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAA 1596
Qy 1659 ACAACCAAGATCTTCTTGGAGGTGATAAGGACATGTGTAGACCCAGATAGAGAAGC 1718
|||||

Db 1597 ACAACCAAGAAATCTTCTTGCAGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAGC 1656
Qy 1719 AAGCGAAGGATTTAGCAATTCCTCGGTGGGTGAACAAAGTTTGAGAGGCTCATCAAAAACC 1778
Db 1657 AAGCGAAGGATTTAGCAATTCCTCGGTGGGTGAACAAAGTTTGAGAGGCTCATCAAAAACC 1716
Qy 1779 AGAAGGAATCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTC 1838
Db 1717 AGAGGGAGTCTCACATTTGTGAGTGTCTGCTCAATCTCAATCTCGTCTC----- 1765
Qy 1839 CTGAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGAGGTC 1898
Db 1766 -----GTCTCTGAAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTC 1815
Qy 1899 CACTCTTTCAATTTTGAAGGCTTTTAACTGAGAAATGAGGCAACTTGTATGTATCGAT 1958
Db 1816 CACTCTTTCAATTTTGAAGGCTTTTAACTGAGAAATGAGGAACTTGTATGTATCGAT 1875
Qy 1959 AATAAGATCAGCGTTTGTACTTACTATCCAAAACTTATCAATAAATAAAAAAGCTTTG 2018
Db 1876 AATAAGATCAGCGTTTGTAACTACTATCCAAAACTTATCAATAAATAAAAAAGCTTTG 1935
Qy 2019 TCGGTTGTTCTCC 2032
Db 1936 TCGGTTGTTCTCC 1949

RESULT 10
AAS08539
ID AAS08539 standard; cDNA; 1952 BP.
AC AAS08539;
DT 23-OCT-2001 (first entry)
XX
DE Anaphylactic antigen Ara h 1, cDNA clone P17.
XX
KW Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
XX allergy; mast cell; basophil; mouse; ss.
XX Mus sp.
XX WO200140264-A2.
XX 07-JUN-2001.
XX 06-DEC-2000; 2000WO-US033124.
XX 06-DEC-1999; 99US-00455294.
XX 23-JUN-2000; 2000US-0213765P.
XX 27-SEP-2000; 2000US-0235797P.
XX (PANA-) PANACEA PHARM LLC.
XX (UYAR-) UNIV ARKANSAS.
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
XX
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals sensitive
PT to antigens, have reduced ability to bind Immunoglobulin E.
XX
XX Disclosure; Fig 9; 100pp; English.
XX
CC The sequence represents anaphylactic antigen Ara h 1, cDNA clone P17. Ara
CC h 1 is an anaphylactic antigen (A), which was used to design antigenic
CC peptides having a reduced ability to bind immunoglobulin E (IgE) as
CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is

QY 1960 ATAAGATCAGCTTTTGTACTTACTATCCAAAACCTTATCAATAATAAAAAAGCTTTGT 2019
|||||
Db 1871 ATAAGATCAGCTTTTGTAACTACTACTATCCAAAACCTTATCAATAATAAAAAAGCTTTGT 1930
|||||
QY 2020 GCCTTGTCTTCTCC 2032
|||||
Db 1931 GCCTTGTCTTCTCC 1943
|||||
RESULT 11
ABS55196
ID ABS55196 standard; DNA; 1251 BP.
XX
AC ABS55196;
XX
DT 17-DEC-2002 (first entry)
XX
DE Glycine max (Soybean) var. Wasesuzunari gene #2.
XX
SW Soybean; Glycinin; atomic coordinate data; processability; soya protein;
KW Wasesuzunari; protein co-ordinate data; gene; ds.
XX
OS Glycine max.
XX
FH Key
FT CDS
FT 1. .1251
FT /*tag= a
FT /partial
FT /product= "Glycine max (Soybean) var. Wasesuzunari
FT protein #2"
FT /note= "this sequence lacks a stop codon"
XX
PN JP2002193996-A.
XX
PD 10-JUL-2002.
XX
PP 21-DEC-2000; 2000JP-00405097.
XX
PR 21-DEC-2000; 2000JP-00405097.
XX
PA (KYOU) UNIV KYOTO.
XX
XX WPI; 2002-685438/74.
DR P-PSDB; ABG71269.
XX
XX Glycinin, beta-conglycinin and proglycinin, their crystal structures,
PT three dimensional coordinates, three dimensional structured and models
PT and their uses.
XX
PS Disclosure; Page 1282-1283; 1298pp; Japanese.
XX
XX The present invention relates to a new Glycinin characterised by the
CC atomic coordinate data fully defined in the specification. The structure
CC can be used for improving processability of soya protein. The present
CC nucleic acid sequence encodes the Glycine max (Soybean) var. Wasesuzunari
CC protein #2, as described in the specification
XX
SQ Sequence 1251 BP; 385 A; 304 C; 265 G; 297 T; 0 U; 0 Other;

Query Match 20.1%; Score 408.4; DB 6; Length 1251;
Best Local Similarity 62.2%; Pred. No. 1.1e-109;
Matches 780; Conservative 0; Mismatches 381; Indels 93; Gaps 5;

QY 562 GAACACCCCTTCTACTT---CCCGTCAGGCGGTTTAGCACCGGTACGGNACCAAAA 618
|||||
Db 24 GAATAACCCCTTCTACTTTAGAGCTCTAACAGCTTCCAACTCTCTTTGAGAACCAAAA 83
|||||
QY 619 CGGTAGGATCCGGGCTCTCGAGAGGTTTGACCAAGGTCAAGGCAAGTTTCAGAAATCTCCA 678
|||||
Db 84 CGGTGCAATTCGTCTCTCCAGAGATTCAACAAACGCTCCCCACAACTTGAGAACCTTCG 143
|||||
QY 679 GAATACCGTATTGTGAGATCGAGGCCAAACCTAAACACTCTTTGTTCTTCCCAAGCACGC 738
|||||

Db 144 AGACTACCGGATGTGCCAGTTTCAGTCAAAAACCCAAACAATCTCTTCTCCCAACATGC 203
QY 739 TGATGCTGATACATCCTTGTATTATCCAGCAAGGGCAAGCCACCGTGACCGTAGCAAAATGG 798
Db 204 TGACGCCGATTTCTCTCTTGTCTTGTCTTAGCGGGAGAGCCATCTTACCTTGGTGAACAA 263
QY 799 CAATAACAGAAAGAGCTTTAATCTTTGACGAGGGCCATGCATTCAGAAATCCCATCCGTTT 858
Db 264 CGACGACAGAGACTCTTACAACCTTTCACCTGGCGATGCCAGAGAAATCCAGCTGGAAC 323
QY 859 CATTTCTTACATCTTTGAACCGGCATGACAAACAGAACTTCAGAGTAGCTAAATCTCCAT 918
Db 324 CACTTACTATTTTGGTTAACTCCTCACGACCCACAGAAATCTCAAAATAATCAAACTTGCAT 383
QY 919 GCCCGTTAAACACACCGGCCAGTTTGTAGGATTTCTTCCGGCGAGCAGCCGAGACCAATC 978
Db 384 ACCCGTCAACAAACCTGGCAGATATGATTTCTTCTTATCTAGACTCAAGCCCAACA 443
QY 979 ATCTTACTTTGCAAGGCTTCAGCAGGAATACGTTGGAGGCCGCTTCAATTCGGAATTCAA 1038
Db 444 GTCTTACTTTGCAAGGCTTCAGCCATAATATCTAGAGACTCTCTTCCATAGCAATTCGA 503
QY 1039 TGAGATACGGAGGCTGCTGTTAGAAAGAAATCCAGAGGTGAGCAAGAGGAGAGAGGCA 1098
Db 504 GGAGATAAACAGGGTTTGTGAGAGGAAGAGGAGCAGAGCGAGCAAGAG----- 555
QY 1099 GAGGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGSAGTAGTAGTCAAAAGTGTCAA 1158
Db 556 -----GGAGTGATCGTGGAACTCTCAAA 578
QY 1159 GGAGCAGCTTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAAAGAAAGGCTCCGAAGA 1218
Db 579 GGAAACAAATTCGGCAACTGAGCAGACGTCGCAAAATCTAGTTCAAGGAAACCAATTTCCCTC 638
QY 1219 AGAGGAGATATCACCAACCCCAATCACTTGAGAGAGCGGAGCCGATCTTCTTCAACAA 1278
Db 639 CGAAGATG-----AACCAATTCAGCTTGAGAGCGGCAACCCCACTCTATTTCCAACAA 689
QY 1279 CTTTGGGAAGTTATTGAGGTGAAGCAGACAAAGAAAGAACCCCGAGCTTCAGGACCTGGA 1338
Db 690 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAACCCACAGCTTCGGGACTTGA 746
QY 1339 CATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGCTCCCACTTCAAACTC 1398
Db 747 TATCTTCTCAGTCTCTGTGATATCAAACGAGGAGCTCTTCTTACACACACTTCAAATTC 806
QY 1399 AAAGGCCATGTTATCTGCTGCTCAACAAAGGAATCGAAACCTTGAACCTCGTGGCTGT 1458
Db 807 AAAGGCCATAGTGATCTAGTGATTAATGAAGGAGATGCAAAACATTGAACCTTGTGGCAT 866
QY 1459 AAGAAAGAGCAACAAACAGAGGGGACGGCGGAAGAGAGGAGGAGCAAGACGAAGAAGA 1518
Db 867 TAAAGAACAAACAAAGAGAGAGAAACAGGAAGAGAACTTTG----- 909
QY 1519 GGAGGGAAGTAAACAGAGAGGTCGTAGGTACACAGCAGGTTTGAAGAGGCGGATGTGTT 1578
Db 910 -----GAAGTCAAAAGGTACAGAGCTGAATTTGTCTGAAGACGATGATTT 953
QY 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACCTCCATCTGTTGG 1638
Db 954 TGTAAATTCAGCAGCTTATCCATTTGTGCTCAACGCTTACCTCAAAACCTCAATTTCTTGC 1013
QY 1639 CTTTGGTATCAACGCTGAAACAAACACAGAAATCTTCTTGTGAGGTGATAAGCAATGT 1698
Db 1014 TTTTGGTATCAATGTGTAGAACACACAGAGGAACCTTCTTGAGGCGGAGAAAGCAATGT 1073
QY 1699 GATAGCCAGATAGAGAAAGCAAGCGAAGGATTTAGCATTCCTCGGTGGGTGAACAAGT 1758
Db 1074 GGTAAAGGCAGATAGAAAAGACAAAGTCAGGAGCTTGGTTCCCTGGGTCTGCACAAGATGT 1133
QY 1759 TGAGAAGCTCATCAAAACCAAGGAATCTCACTTTGTGAGTGTGCTGCTCTCA 1812
|||||

Db 1134 TGAGAGGCTATTAAAGAGCAGAGGAATCCTACTTTGTTGATGCTCAGCCTCA 1187

RESULT 12

ACC49562

ID ACC49562 standard; cDNA; 1254 BP.

XX AC ACC49562;

XX AC ACC49562;

XX 01-JUL-2003 (first entry)

XX DE

XX Mature beta-conglycinin beta subunit encoding cDNA SEQ ID NO:15.

XX KW

XX Glycine max; soybean; 7S-beta-conglycinin beta subunit; plant; food; nutritional supplementation; animal feed; nutritive value; malnutrition; health; gene; ss.

XX KW

XX Glycine max.

XX OS

XX Key

XX Location/Qualifiers

XX 1..1254

XX /*tag= a

XX /product= "mature beta-conglycinin beta subunit"

XX FT

XX MO2003025153-A2.

XX PN

XX 27-MAR-2003.

XX PD

XX 17-SEP-2002; 2002MO-US030373.

XX PP

XX 17-SEP-2001; 2001US-0322461P.

XX PR

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX PA

XX Rapp WD, Peng J, Nadig G, Venkatesh T;

XX PI

XX WPI: 2003-333163/31.

XX DR

XX P-PSDB; ABP97239.

XX XX

XX Novel modified polypeptide useful in nutritional supplementation for animals and as human food, comprises a substitution of essential amino acids such as lysine, threonine and tryptophan into an unmodified polypeptide.

XX PS

XX Example 6; Page 91; 107pp; English.

XX CC

The present invention describes a modified polypeptide (I) comprising a substitution of one or more essential amino acids selected from isoleucine, lysine, methionine, threonine and tryptophan into an unmodified polypeptide having a sequence (SI) of 439 amino acids (see ABP97235) of the Glycine max 7S-beta-conglycinin beta subunit, where the modified polypeptide is capable of accumulating in a seed. Also described: (i) a recombinant nucleic acid molecule (II) encoding (I), preferably a modified beta-conglycinin polypeptide, where the modified beta-conglycinin polypeptide is capable of accumulating in a cell; (2) a cell (III) comprising (II); (3) a transgenic plant (IV) comprising (III); (4) a seed (V) from (IV); and (5) an animal feed comprising (V). (I) is useful in nutritional supplementation for animals, including humans. (II), (IV) and (V) are useful as human food. (I) significantly improves the nutritive value of animal feed, leading to improved quality and quantity of commercial animal products, and also significantly improves the nutritive value of human food, leading to a decreased incidence of malnutrition and associated health problem and improves the overall growth and development of infants and children. The present sequence encodes a mature beta-conglycinin beta subunit, which is used in the exemplification of the present invention

XX SQ

XX Sequence 1254 BP; 385 A; 305 C; 268 G; 296 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 20.0%; Score 406.8; DB 9; Length 1254;

XX Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;

QY 562 GAACACCCCTTTCTACTT---CCCGTCAAGCGGTTTAGCACCGCTACGGGACCAAAA 618
Db 24 GAATAACCCCTTTCTACTTTGAGAGCTCTAACAGCTTCCAAACTCTCTTTGAGAACCAAAA 83
QY 619 CGGTAGGATCCGGGTCTCTCGAGAGGTTTGACCAAGGTCAAGGAGTTCAGATCTCTCCA 678
Db 84 CGGTGCGATTTCGTCTCTCTCCAGAGATTCAACAAACGCTCCCAACACTTGAGAACCTTCG 143
QY 679 GAATCACCGTATTGTGTCAGATCGAGGCCAAACCTTAAACACTCTCTGTTTCTCCCAAGCAGC 738
Db 144 AGACTACCGGATTGTCCAGTTTCAGTCAAAACCCCAACACAATCTCTTCCCTCCCAACATGC 203
QY 739 TGATGCTGATAAATCTCTGTTATCCAGCAAGGCGAGCCACCGTGACCTGACCTAGCAATGG 798
Db 204 TGACGCGGATTCTCTCTCTTGTCTTACGGGAGAGCATATCTTACCTTGGTGGAAACA 263
QY 799 CAATAACAGAAAGAGCTTTTAATCTTGACGAGGCGCATGCACTCAGAAATCCCATCCGGTTT 858
Db 264 CGACGACAGAGACTCTTACAACTTCACTCTGCGATGCCAGAGAAATCCAGCTGGAAC 323
QY 859 CATTTCCTCATCTTGAAACCGCATGACAAACCAAGCTCAGAGTAGCTAAATCTCCAT 918
Db 324 CACTTACTATTGTGTTAAACCTCAGCAACCAAGAAATCTCAAAATATCAAACTTGGCAT 383
QY 919 GCCCGTTAACACACCGCGCAGTTTGAGGATTTCTTCCCGCGAGGAGCGAGACCAATC 978
Db 384 ACCCGTCAACAAACCTGGCAGATATGATGATTTCTTCTTATCTAGCACTCAAGGCCAACA 443
QY 979 ATCTTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGGCGGCTTCAATCGGAATTCAA 1038
Db 444 GTCTTACTTGCAGGGCTTCAGCCATATATTTCTAGAGACCTCTTCTCATAGCGAATTCGA 503
QY 1039 TGAGATACGGAGGGTGTGTTTGAAGAGAAATGACGAGGTGACAAAGAGAGAGAGGCA 1098
Db 504 GGAGATAAACACGGGTTTGTGGAGAGGAAGAGGAGCAGAGCAGCAAG-----555
QY 1099 GAGCGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGTGTATAGTCAAAAGTGCAAA 1158
Db 556 -----GGAGTGATCGTGAATCTCTCAA 578
QY 1159 GGAGCAGCTTTGAAGAACTTACTAAGCAGCTAAATCGCTCAAGAAAGAGCTCCGAAGA 1218
Db 579 GGAACAAATTCGCAACTGAGCAGAGCTGCCAATCTAGTTCAAGGAAACCAATTTCTCTC 638
QY 1219 AGAGGAGATATCACCACCCAACTTGAAGAAAGGCGAGCCCGATCTTTCTAACAA 1278
Db 639 CGAAGATG-----AACCAATTCACTTGAAGCGCAACCCCACTTATTTCCAACA 689
QY 1279 CTTTGGGAAGTTATTTGAGTGAAGCCAGACAGAAAGAACCCCGAGCTTCAGACCTGGA 1338
Db 690 CTTTGGAAAGTTCTTTGAGATCACCCTG---AGAAACCCCAAGCTTCGGGACTTGG 746
QY 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCCACTTCAACTC 1398
Db 747 TATCTTCTCAGTCTGTGGATATCAAGAGAGCTCTTCTTCTACCACTTCAATTC 806
QY 1399 AAAGGCCATGTTTATCGTCTCGTCAACAAAGGAACCTTGAACCTTCGCGGTGT 1458
Db 807 AAAGGCCATGATGATCTAGTGATTAAAGAGGAGATGCAAACTTGAACCTTGTGGCAT 866
QY 1459 AGAAAGAGCAACACAGAGGGGAGCGGGGAGAAAGAGGAGGAGCAGACGAGAGAGA 1518
Db 867 TAAAGAACCAACACAGAGCAGAAACAGGAAGGAAGCACTTTG-----909
QY 1519 GGAGGGAAGTAAACAGAGAGTGTGTAGGTACACAGCGAGGTTCGAAGGAGCGATGTGT 1578
Db 910 -----GAAGTGAAGGTACAGAGCTGATTTGTTGTAAGCAATGTATT 953
QY 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACCGTTCTCCGAATCTCCATCTGCTGG 1638
Db 954 TGTAAATTCAGCAGCTTATCCATTTGTGTCGTCACGCTACCTCAACACCTCAATTTCTTGC 1013
QY 1639 CTTCCGTTATCAACGCTGAAACCAACCAAGAAATCTTCTTGGAGGTGATGAAGCAATGT 1698

Db 1014 TTTTGGTATCAATGCTCAGAACCAACAGAGGAACTTCTTTCAGGCGAGAAAGACAATGT 1073
Qy 1699 GATAGACCAGATAGAGCAAGCAAGAGGATTTAGCATTTCCCTGGGTGCGGTGAACAAGT 1758
Db 1074 GGTAAAGGCAGATAGAAAGACAAGTGCAGGAGCTTGGCTTCCCTGGGTGCTGCAAGAATGT 1133
Qy 1759 TGAGAGCTCATCAAAACAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
Db 1134 TGAGAGGCTATTAAAGAGCAGAGGGAATCTACTTTGTGTGCTCAGCCTCA 1187

RESULT 13
ACC49561
ID ACC49561 standard; cDNA; 1278 BP.
XX ACC49561;
AC ACC49561;
XX 01-JUL-2003 (first entry)
DT FLAG-tagged beta-conglycinin beta subunit encoding cDNA SEQ ID NO:13.
DE
XX
XX Glycine max; soybean; 7S-beta-conglycinin beta subunit; plant; food;
KW nutritional supplementation; animal feed; nutritive value; malnutrition;
KW health; gene; ss.
XX
XX Glycine max.
OS
OS Synthetic.

XX Key Location/Qualifiers
FH 1. .1278
FT CDS /*tag= a
FT /product= "FLAG-tagged beta-conglycinin beta subunit"
FT

PN W02003025153-A2.

XX 27-MAR-2003.

XX 17-SEP-2002; 2002WO-US030373.

XX 17-SEP-2001; 2001US-0322461P.

XX (MONS) MONSANTO TECHNOLOGY LLC.

XX Rapp WD, Peng J, Nadig G, Venkatesh T;

XX WPI; 2003-333163/31.

XX P-PSDB; ABP97238.

XX Novel modified polypeptide useful in nutritional supplementation for
PT animals and as human food, comprises a substitution of essential amino
PT acids such as lysine, threonine and tryptophan into an unmodified
PT polypeptide.

XX Example 6; Page 88-89; 107pp; English.

PS The present invention describes a modified polypeptide (I) comprising a
CC substitution of one or more essential amino acids selected from
CC isoleucine, lysine, methionine, threonine and tryptophan into an
CC unmodified polypeptide having a sequence (S1) of 439 amino acids (see
CC ABP97235) of the Glycine max 7S-beta-conglycinin beta subunit, where the
CC modified polypeptide is capable of accumulating in a seed. Also
CC described: (1) a recombinant nucleic acid molecule (II) encoding (I),
CC preferably a modified beta-conglycinin polypeptide, where the modified
CC beta-conglycinin polypeptide is capable of accumulating in a cell; (2) a
CC cell (III) comprising (II); (3) a transgenic plant (IV) comprising (III);
CC (4) a seed (V) from (IV); and (5) an animal feed comprising (V). (I) is
CC useful in nutritional supplementation for animals, including humans. (I),
CC (IV) and (V) are useful as human food. (I) significantly improves the
CC nutritive value of animal feed, leading to improved quality and quantity
CC of commercial animal products, and also significantly improves the
CC nutritive value of human food, leading to a decreased incidence of
CC malnutrition and associated health problem and improves the overall

QY 1519 GGAGGAGTAAACAGAGAGTCCGTAGGTACACAGCAGAGTTGAAGGAGCGATGTGT 1578
 Db 934 -----GAAGTGCAAGGTACAGAGCTGAATTGTCTGAAGACGATGATT 977
 QY 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAAGCTTCCTCCGAACTCCCATCTGCTTGG 1638
 Db 978 TGTAAATCCAGCAGCTTATCCATTGTCTGTCAGCGTACCTCAACCTCAATTCCTTGC 1037
 QY 1639 CTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATAGGACAATGT 1698
 Db 1038 TTTTGGTATCAATGCTGAGAAACCAACAGAGAACTTCTTCAGGCGAGAAAGACAATGT 1097
 QY 1699 GATAGCCAGATAGAGAGCAGGAGGATTTAGCATTCCTCGGTCCGCTGACAAAT 1758
 Db 1098 GGTAGGCGATAGAGAGCAAGTGCAGAGCTTGCCTTCCTGCGTCTGCACAAATGT 1157
 QY 1759 TGAGAAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
 Db 1158 TGAGAGGCTATTAAAGAGCAGAGGGAATCCTACTTTGTTGATGCTCAGCCTCA 1211

RESULT 14

AAV17564
 ID AAV17564 standard; cDNA; 1320 BP.

AC AAV17564;

XX 10-JUN-1998 (first entry)

XX Coding sequence for the beta subunit of beta-conglycinin.

XX Beta-conglycinin; soybean seed protein; transgenic plant;

KW seed storage protein profile; ss.

XX Glycine max.

XX W09747731-A2.

XX 18-DEC-1997.

XX 10-JUN-1997; 97WO-US009743.

XX 14-JUN-1996; 96US-0019940P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Kinney AJ, Fader GW;

XX WPI; 1998-052298/05.

XX Suppression of specific classes of soybean seed protein genes - useful to
 change seed storage protein profiles of transgenic plants.

XX Disclosure; Page 31-32; 58pp; English.

XX This sequence represents the coding sequence for the beta subunit of the
 soybean seed protein beta-conglycinin. The method of the invention is for
 reducing the quantity of a soybean seed storage protein (A), such as beta
 -conglycinin, in soybeans. It comprises: (a) constructing a chimeric gene
 comprising: (i) a nucleic acid fragment encoding a promoter that is
 functional in the cells of soybean seeds; (ii) a nucleic acid fragment
 encoding all or a portion of (A) placed in sense or antisense orientation
 relative to the promoter of (i); and (iii) a transcriptional termination
 region; (b) creating a transgenic soybean cell by introducing into a
 soybean cell the chimeric gene of (a); and (c) growing the transgenic
 soybean cells of (b) under conditions that result in expression of the
 chimeric gene of (a); where the quantity of one or more members of a
 class of (A) subunits is reduced when compared to soybeans not containing
 the chimeric gene of (a). The method is used to construct transgenic
 soybean lines where the expression of genes encoding (A) are modulated to
 effect a change in seed storage protein profile of transgenic plants.
 XX Modification of the seed storage protein profile can result in the

CC production of novel soy protein products with unique and valuable
 CC functional characteristics
 XX
 SQ Sequence 1320 BP; 394 A; 315 C; 285 G; 326 T; 0 U; 0 Other;
 Query Match 20.0%; Score 406.8; DB 2; Length 1320;
 Best Local Similarity 62.1%; Pred. No. 3.5e-109;
 Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
 QY 562 GAACAACCCCTTTCTACTT---CCCGTCAAGCGGTTTACACCCGCTACGGGAACCAAAA 618
 Db 90 GAATAACCCCTTTCTACTTTAGAAAGCTCTAAACAGCTTCCAAACTCTCTTTGAGAACCAAAA 149
 QY 619 CGGTAGGATCCGGGTCTCGAGAGTTTGACCAAGGTCAAGGAGTTTCAGATCTCCA 678
 Db 150 CGTTCGCAATTCGTCTCTCCAGAGATTCAACAAACGCTCCCAACCTTGAGAACCTTCG 209
 QY 679 GAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTGTCTTCCCAAGCACGC 738
 Db 210 AGACTACCGGATTGTCCAGTTTCAGTCAAAACCCACACAACTCTTCTCCCCACCATGC 269
 QY 739 TGATGTGATAACATCTTGTATTCCAGCAAGGCAAGCCACCGTGAACCTGAGCAAAATGG 798
 Db 270 TGACGCGGATTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 329
 QY 799 CAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCACTCAGATCTCCATCCGTTT 858
 Db 330 CGACGACAGAGACTCTCTACAACTTCAACCTTCCCTGCGGATGCCAGAGAAATCCAGCTGAAC 389
 QY 859 CATTTCTACATCTTGAAACCGCATGACAAACCAAGAACTCAGAGTAGTAAATCTCCAT 918
 Db 390 CATTACTATTGGTTTAACTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 449
 QY 919 GCCCGTTAACACACCCCGGCGAGTTTGAGGATTTCTTCCGCGGAGGAGGAGGAGGAGGAGGAG 978
 Db 450 ACCCGTCAACAAACCTGGCAGATATGATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 509
 QY 979 ATCTACTTTCAGGCGTTTACGAGGAAATAGTTTGGAGGCGGCTTCAATCGGGAATTCAA 1038
 Db 510 GTCTACTTTCAGGCGTTTACGAGGAAATAGTTTGGAGGCGGCTTCAATCGGGAATTCGA 569
 QY 1039 TGAGATACGGAGGCTGCTGTAGAGAGAAATGCAAGGAGGTGAGCAAGAGGAGGAGGAGGAGGAG 1098
 Db 570 GGAGATTAACAGGCTTTTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
 QY 1099 GAGCGATGGAGTACTCGGAGTAGTGAGAAATAAGAGGAGTAGTAGTCAAGAGTGTCAA 1158
 Db 622 -----GGAGTGTCTGTGGAACCTCTCAA 644
 QY 1159 GGAGCAGCTTGAAGAACTTACTAAGCAGCGTAAATCGTCTCAAGAAAGGCTCCGGAAGA 1218
 Db 645 GGAACAAATTCGGCAACTGAGCAGAGCGTGCATAATCTAGTTCAAGGAAACCAATTTCCCTC 704
 QY 1219 AGAGGAGATATCACCAACCAATCACTTGAGAGAGGCGGAGCGGAGTCTTCTTCAACA 1278
 Db 705 CGAAGATG-----AACCATTAATTTGAGAGCGGCAACCCCATCTATTTCCAACAA 755
 QY 1279 CTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAACCCCGAGCTTCAGGACCTGGA 1338
 Db 756 CTTTGGGAAGTTCTTTGAGATCACCCCTG---AGAAACCCACAGCTTCGGGACTTGA 812
 QY 1339 CATGATGCTCACCTGTGTAGAGATCAAGAGAGCTTTGATGCTCTCCACACTTCAACTC 1398
 Db 813 TATCTTCTCAGTTCTGTGGATATCAACGAAGGAGCTCTTCTTCTACACACTTCAATTC 872
 QY 1399 AAAGGCCATGTTTATCGTCTGTCAACAAAGAACTGGAAACCTTGAACCTTCTGCTGT 1458
 Db 873 AAAGGCCATGATGATAGTGTGATTAATGAAGGAGATGCAAACTTGAACCTTGTGGCAT 932
 QY 1459 AAGAAAGAGCAACACAGAGGCGGCGGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1518
 Db 933 TAAAGAACACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 975

Qy	1339	CATGATGCTCACTGTGTAGAGATCAAGAAGAGGAGCTTTGATGCTCCACACTTCAACTC	1398
Db	813	TATCTTCTCTAGTTCTGTGGATATCAACGAGGAGCTCTTCTTCTACACACTTCAATTC	872
Qy	1399	AAAGGCCATGGTTATCGTCGTCGTCACAAGAAGAACTGGAAACCTTGAACCTCGTGCGCTGT	1458
Db	873	AAAGGCCATAGTGATACTAGTGATTATGAAGGAGATGCAAAACATTGACTTTGTTGGCAT	932
Qy	1459	AAGAAAAGAGCAACAA CAGAGGGGACGGCGGGAAAGAGAAGGAGCAAGAA CCGAAGAAGA	1518
Db	933	TAAAGAAACAA CAA CAGAGAGCAGAAACAGGAAGAGGAACCTTTG-----	975
Qy	1519	GGAGGGAATACAGAGAGGTGCGTAGGTACACAGCGAGGTGGAAGAAAGCGCATGTGTT	1578
Db	976	-----GAAGTGCAAAAGGTACAGAGCTGAATTGCTGAGACGAGTATTT	1019
Qy	1579	CATCATCCAGCAGCTCATCCAGTAGGCATCAACGCTTCTCTCCGAACTCCATCTGCTGG	1638
Db	1020	TGTAATTTCCAGCAGCTTATCCATTGTTCGTCAACGCTACCTCAAAACCTCAATTTCCTTGC	1079
Qy	1639	CTTCGGTATCAACGCTGAAACACACACAGNATCTTCTTCGAGGTGATTAAGACAAATGT	1698
Db	1080	TTTGTGGTATCAATGCTGAGAACCAACAGAGGAACCTTCTTCGAGGCGAGAAAGACAATGT	1139
Qy	1699	GATAGACAGATAGAGAAGCAAGCGAAGGATTTTAGCATTCCTCTGGGTCGGTCAACAAGT	1758
Db	1140	GGTAAGGCAGATAGAAAGACAAAGTGCAGGAGCTTGCCTTCCTGGGTCTGCACAAAGTGT	1199
Qy	1759	TGAGAAGCTCATCAAAACCAGAAAGGAATCTCACTTTGTGAGTGCTCGTCTCTCA	1812
Db	1200	TGAGAGGCTATTAAAGAAAGCAGAGGGAATCTACTTTGTTGATGCTCAGGCTCA	1253

Search completed: August 23, 2005, 21:26:04
Job time : 773.871 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 334.307 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aatacatatattatcatc.....cgtttgctggtttcttc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	4	US-09-106-872A-21
2	2032	100.0	2032	4	US-09-191-593-5
3	2032	100.0	2041	4	US-09-106-872A-23
4	1930	95.0	1930	4	US-09-106-872A-3
5	1752.4	86.2	1949	4	US-09-106-872A-19
6	1752.4	86.2	1949	4	US-09-191-593-4
7	750	36.9	750	4	US-09-106-872A-18
8	406.8	20.0	1320	3	US-09-108-010B-3
9	406.8	20.0	1320	4	US-09-758-552-3
10	406.8	20.0	1320	4	US-10-684-651-3
11	404	19.9	1818	3	US-09-108-010B-1
12	404	19.9	1818	4	US-09-758-552-1
13	404	19.9	1818	4	US-10-684-651-1
14	324.2	16.0	1920	3	US-09-108-010B-2
15	324.2	16.0	1920	4	US-09-758-552-2
16	324.2	16.0	1920	4	US-10-684-651-2
17	138.8	6.8	1867	1	US-07-955-905A-1
18	98.8	4.9	1924	3	US-09-424-283-5
19	71.6	3.5	7218	1	US-08-232-463-14
20	49.4	2.4	1068	4	US-09-710-794-3
21	46.6	2.3	74177	4	US-09-949-016-11988
22	46.6	2.3	74177	4	US-09-949-016-17388
23	46	2.3	1590	3	US-09-323-195A-7
24	45.8	2.3	300402	4	US-09-949-016-13632
25	44.4	2.2	499	4	US-10-235-618B-16
26	44.2	2.2	61664	4	US-09-949-016-13308
27	43.8	2.2	767677	4	US-09-949-016-12147

28	43.8	2.2	767677	4	US-09-949-016-17361	Sequence 17361, A
29	43.4	2.1	1318	4	US-10-235-618B-1	Sequence 1, Appl
30	43.4	2.1	1318	4	US-10-235-618B-11	Sequence 11, Appl
31	43.2	2.1	289	3	US-09-007-005-17	Sequence 17, Appl
32	43.2	2.1	289	3	US-09-244-796-17	Sequence 17, Appl
C 33	42	2.1	150394	4	US-09-949-016-13042	Sequence 13042, A
C 34	41.8	2.1	28806	4	US-09-949-016-13217	Sequence 13217, A
35	41	2.0	73519	4	US-09-949-016-16344	Sequence 16344, A
36	41	2.0	105919	4	US-09-949-016-11769	Sequence 11769, A
C 37	40.8	2.0	50383	4	US-09-949-016-17600	Sequence 17600, A
C 38	40.8	2.0	234884	4	US-09-949-016-16420	Sequence 16420, A
C 39	40.6	2.0	46823	4	US-09-949-016-12723	Sequence 12723, A
C 40	40.6	2.0	46940	4	US-09-949-016-16252	Sequence 16252, A
C 41	40.6	2.0	69813	4	US-09-949-016-12455	Sequence 12455, A
C 42	40.6	2.0	69813	4	US-09-949-016-13906	Sequence 13906, A
C 43	40.6	2.0	69833	4	US-09-949-016-12861	Sequence 12861, A
C 44	40.6	2.0	69833	4	US-09-949-016-12861	Sequence 12861, A
C 45	40.4	2.0	601	4	US-09-949-016-167942	Sequence 167942, A

ALIGNMENTS

RESULT 1

US-09-106-872A-21
; Sequence 21, Application US/09106872A

; Patent No. 6486311

; GENERAL INFORMATION:

; APPLICANT: Burks Jr., A. Wesley

; APPLICANT: Stanley, J. Steven

; APPLICANT: Cockrell, Gael

; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.

; APPLICANT: Helm, Ricki M.

; APPLICANT: Bannon, Gary A.

; TITLE OF INVENTION: Peanut Allergens and Methods

; FILE REFERENCE: HS 103 CIP

; CURRENT APPLICATION NUMBER: US/09/106,872A

; CURRENT FILING DATE: 1999-06-29

; PRIOR APPLICATION NUMBER: PCT/US96/15222

; PRIOR FILING DATE: 1996-09-23

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 21

; LENGTH: 2032

; TYPE: DNA

; ORGANISM: Arachis hypogaea

US-09-106-872A-21

Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60
Db	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60
Qy	61	GGTTTCTCCACTGATGCTGTGTAGGGATCTTGTCTCGCTTCAGTTTCTGCAAGCA	120
Db	61	GGTTTCTCCACTGATGCTGTGTAGGGATCTTGTCTCGCTTCAGTTTCTGCAAGCA	120
Qy	121	TGCAAGTATCATCTTACCAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA	180
Db	121	TGCAAGTATCATCTTACCAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTTCCA	180
Qy	181	GAGTTCTCAACAGAACCCGATGCTTGAAGCAAAAGGAGTCTCGCTGCACCA	240
Db	181	GAGTTCTCAACAGAACCCGATGCTTGAAGCAAAAGGAGTCTCGCTGCACCA	240
Qy	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA	300
Db	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA	300

Qy 301 ACCTTCCCTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Db 301 ACCTTCCCTCCAGGGAGCGGACACGCTGGCCGCCAACCCGGAGACTACGATGATACCG 360
Qy 361 CCCTCAACCCCGAAG 420
Db 361 CCCTCAACCCCGAAG 420
Qy 421 AAG 480
Db 421 AAG 480
Qy 481 ACAG 540
Db 481 ACAG 540
Qy 541 TGTGAGGAG 600
Db 541 TGTGAGGAG 600
Qy 601 CCGCTACCGGAAACCAAAACCGTGGATCCGGGCTCTGACAGAGTTTGAACCAAAAGGTTCAAG 660
Db 601 CCGCTACCGGAAACCAAAACCGTGGATCCGGGCTCTGACAGAGTTTGAACCAAAAGGTTCAAG 660
Qy 661 GCAGTTTTCAGAACTCCAGAACTCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCT 720
Db 661 GCAGTTTTCAGAACTCCAGAACTCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCT 720
Qy 721 TGTCTTCCAGACGCTGATGCTGATTAACATCTCTGTTATCCAGCAGAGGCGCAAGCCAC 780
Db 721 TGTCTTCCAGACGCTGATGCTGATTAACATCTCTGTTATCCAGCAGAGGCGCAAGCCAC 780
Qy 781 CGTGACGCTAGCAAAATGCGCAATACAGAAAGAGCTTTAACTTTGACAGAGGCGCATGCAC 840
Db 781 CGTGACGCTAGCAAAATGCGCAATACAGAAAGAGCTTTAACTTTGACAGAGGCGCATGCAC 840
Qy 841 CAGAACTCCATCCGGTTTCAATTTCTTACATCTTTGAACCGGCATGACAAACCAAGACCTCAG 900
Db 841 CAGAACTCCATCCGGTTTCAATTTCTTACATCTTTGAACCGGCATGACAAACCAAGACCTCAG 900
Qy 901 AGTAGCTTAAATCTCCATGCGGTTTAAACACACCGGCGAGTTTGAGGATTTCTTCCGGC 960
Db 901 AGTAGCTTAAATCTCCATGCGGTTTAAACACACCGGCGAGTTTGAGGATTTCTTCCGGC 960
Qy 961 GAGCAGCGGAGACCAATCTCTACTTTCAGGGCTTCAGCAGGAATACGTTGGAGGCGC 1020
Db 961 GAGCAGCGGAGACCAATCTCTACTTTCAGGGCTTCAGCAGGAATACGTTGGAGGCGC 1020
Qy 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGCTGTTAGAGAGAAATCGAGAGGTGA 1080
Db 1021 CTTCAATGCGGAATTCATGAGATACGAGGGTGCTGTTAGAGAGAAATCGAGAGGTGA 1080
Qy 1081 GCAGAGGAGAGAGGCGAGAGGCGATGAGTACTCGGAGTAGTGAGAACATCAAGAGGT 1140
Db 1081 GCAGAGGAGAGAGGCGAGAGGCGATGAGTACTCGGAGTAGTGAGAACATCAAGAGGT 1140
Qy 1141 GATAGTCAAAAGTTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCCTAAATCCGCTC 1200
Db 1141 GATAGTCAAAAGTTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCCTAAATCCGCTC 1200
Qy 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATCACAAACCCCAATCAACTTTGAGAGAGCGA 1260
Db 1201 AAAGAAAGGCTCCGAGAGAGAGGAGATATCACAAACCCCAATCAACTTTGAGAGAGCGA 1260
Qy 1261 GCCCGATCTTTTCAACAACTTTGGGAGTTATTGAGGTGAAGCCAGACAGAGAACCC 1320
Db 1261 GCCCGATCTTTTCAACAACTTTGGGAGTTATTGAGGTGAAGCCAGACAGAGAACCC 1320
Qy 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGAGCTTTGAT 1380
Qy 1381 GCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCTCGTCAACAAAGGAACTGGAAA 1440

Db 1381 GCTCCCACTTCAACTCAAAAGGCCATGGTTATCGTCTCGTCAACAAAGGAACTGGAAA 1440
Qy 1441 CTTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAAACAGAGGGGACCGCGGAGAGAGGA 1500
Db 1441 CTTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAAACAGAGGGGACCGCGGAGAGAGGA 1500
Qy 1501 GGAACGAAGACGAAGTT 1560
Db 1501 GGAACGAAGACGAAGTT 1560
Qy 1561 GAAGGAAGGCGATGTTTCATCATGCTCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
Db 1561 GAAGGAAGGCGATGTTTCATCATGCTCAGCAGCTCATCCAGTAGCCATCAACGCTTCTC 1620
Qy 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAACTTCTCTTGC 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAAGAACTTCTCTTGC 1680
Qy 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGAGATTTAGCATTTCC 1740
Db 1681 AGGTGATAAGGACAATGTGATAGACAGATAGAGAGCAAGCAAGAGATTTAGCATTTCC 1740
Qy 1741 TGGGTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGAAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCGGGTGAACAAAGTTGAGAGCTCATCAAAACCAAGAGAAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGCTCGCTCGTGAAGAGAGTCTCTCAGAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCGCTCGCTCGTGAAGAGAGTCTCTCAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Db 1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Qy 1921 TTTTAACGAGAAATGAGGCAACTTGTATGATCGATAATAGATCAACGCTTTTGTACT 1980
Db 1921 TTTTAACGAGAAATGAGGCAACTTGTATGATCGATAATAGATCAACGCTTTTGTACT 1980
Qy 1981 CTACTATCCAAACCTTATCAATAAATAAAACGTTTGTGCGTTTCTTCC 2032
Db 1981 CTACTATCCAAACCTTATCAATAAATAAAACGTTTGTGCGTTTCTTCC 2032

RESULT 2
US-09-191-593-5
; Sequence 5, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

1081 GCAAGAGGAGAGGGGACAGAGGCGATCGAGTACTCGGAGTAGTGAGAACAAATGAAGAGT 1140
1141 GATAGTCAAAAGTGTCAAGAGGACCGTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTCC 1200
1141 GATAGTCAAAAGTGTCAAGAGGACCGTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTCC 1200
1201 AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTGAAGAGGCGA 1260
1201 AAAGAAAGGCTCCGAAGAGGAGGAGATATCACCAACCCCAATCAACTTGAAGAGGCGA 1260
1261 GCCGATCTTTTCAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGCAAGAAAGAACCC 1320
1261 GCCGATCTTTTCAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGCAAGAAAGAACCC 1320
1321 CCAGCTTCAGGACTCGACATGATGCTCACCTGTGTAGAGTCAAGAGGAGCTTTGAT 1380
1321 CCAGCTTCAGGACTCGACATGATGCTCACCTGTGTAGAGTCAAGAGGAGCTTTGAT 1380
1381 GCTCCACACACTTCAACTCAAAAGCCATGGTTATCGTCTGCTCAACAAAGGAACTCGAAA 1440
1381 GCTCCACACACTTCAACTCAAAAGCCATGGTTATCGTCTGCTCAACAAAGGAACTCGAAA 1440
1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
1501 GACAGAGACGAAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTACACAGCGAGGTT 1560
1501 GACAGAGACGAAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTACACAGCGAGGTT 1560
1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCCCTC 1620
1561 GAAGGAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCCCTC 1620
1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAACTTCTCCTTGC 1680
1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAACTTCTCCTTGC 1680
1681 AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1740
1681 AGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAGGAGGAGGAGGAGGAGGAGGAGG 1740
1741 TGGGTGCGGTGCAACAGTTCGAGAGCTCATCAAAACCAAGAGGAGTCTCACCTTTGTGAG 1800
1741 TGGGTGCGGTGCAACAGTTCGAGAGCTCATCAAAACCAAGAGGAGTCTCACCTTTGTGAG 1800
1801 TGCTCGTCTCAATCTCAATCTCCGTCGTCCTCGTCTCTCGTGAAGAGAGTCTCTCGAGAA 1860
1801 TGCTCGTCTCAATCTCAATCTCCGTCGTCCTCGTCTCTCGTGAAGAGAGTCTCTCGAGAA 1860
1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1861 AGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
1921 TTTTAACTGAGATGAGAGGCAACTTGTATGTATCGATATAGATCAAGATCACGCTTTGTACT 1980
1921 TTTTAACTGAGATGAGAGGCAACTTGTATGTATCGATATAGATCAAGATCACGCTTTGTACT 1980
1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGCTTTGTGCGTTGTTTCTCC 2032
1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGCTTTGTGCGTTGTTTCTCC 2032

RESULT 3

US-09-106-872A-23
; Sequence 23, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.

; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricky M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 2041
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-106-872A-23

Query Match 100.0%; Score 2032; DB 4; Length 2041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGGAG 60
Qy 61 GGTTCCTCCACTGATGCTGTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA 120
Db 61 GGTTCCTCCACTGATGCTGTGCTAGGGATCCTTGTCTCTGGCTTCAGTTTCTGCAACGCA 120
Qy 121 TGCAAAGTCATCACTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCCTCCA 180
Db 121 TGCAAAGTCATCACTTACAGAGAAACAGAGAAACCCCTGCGCCAGAGGTGCCTCCA 180
Qy 181 GAGTGTCTCAACAGGACCGGATGACTTGAAGCAAAAGGATCGAGCTTCGCTGCACCAA 240
Db 181 GAGTGTCTCAACAGGACCGGATGACTTGAAGCAAAAGGATCGAGCTTCGCTGCACCAA 240
Qy 241 GCTCGAGTATGATTCCTCGTTGTCTATGATCCTCGAGGACACACTGGCAACCAACCA 300
Db 241 GCTCGAGTATGATTCCTCGTTGTCTATGATCCTCGAGGACACACTGGCAACCAACCA 300
Qy 301 AGTTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTACGATGATGACCG 360
Db 301 AGTTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTACGATGATGACCG 360
Qy 361 CCGTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACACAGCTTGGACCCGAGGAGCGTGA 420
Db 361 CCGTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACACAGCTTGGACCCGAGGAGCGTGA 420
Qy 421 AAGAGAAGAGAGTGGAGACAACCAAGAGAAGATTGGAGCGGACCAAGTCAATCAGCAGCC 480
Db 421 AAGAGAAGAGAGTGGAGACAACCAAGAGAAGATTGGAGCGGACCAAGTCAATCAGCAGCC 480
Qy 481 ACGGAAAATAAGCCCGAAG 540
Db 481 ACGGAAAATAAGCCCGAAG 540
Qy 541 TGTGAGGAG 600
Db 541 TGTGAGGAG 600
Qy 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAG 660
Db 601 CCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCGAGAGGTTTGACCAAGGTCAG 660
Qy 661 GCAGTTTCAGAAATCTCCAGAAATCAGGATCGAGGCGCAACCTTAACACTCT 720
Db 661 GCAGTTTCAGAAATCTCCAGAAATCAGGATCGAGGCGCAACCTTAACACTCT 720
Qy 721 TGTTCCTTCCCAAGCAGCTGATGATCAATCAATCTTGTATCCAGCAAGGGCAAGCCAC 780
Db 721 TGTTCCTTCCCAAGCAGCTGATGATCAATCAATCTTGTATCCAGCAAGGGCAAGCCAC 780

SOFTWARE: Wordperfect 6.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/191,593
FILING DATE: 13 NOVEMBER 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/717,933
FILING DATE: 23 SEPTEMBER 1996
APPLICATION NUMBER: US 07/998,377
FILING DATE: 30 DECEMBER 1992
APPLICATION NUMBER: US 08/158,704
FILING DATE: 29 NOVEMBER 1993
APPLICATION NUMBER: US 60/009,455
FILING DATE: 29 DECEMBER 1995
APPLICATION NUMBER: US 08/610,424
FILING DATE: 04 MARCH 1996
ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1949 bases
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: identified as Ara h I Alpha P17
HYPOTHEICAL: No
ANTI-SENSE: No
FRAGMENT TYPE: No. 6835824 applicable
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE: Clone P17
DEVELOPMENTAL STAGE: Seed
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE: Seed mRNA, cDNA library
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY: Florunner seed cDNA expression
LIBRARY: library in Uni-ZAP XR vector
CLONE: P17
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1847
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:

US-09-191-593-4

Query Match 86.2%; Score 1752.4; DB 4; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;
QY 48 CAATGAGAGGAGGGTTTCTCCACATGATGCTGTCTAGGATTCCTTGTCTGGCTTCAG 107
DB 1 CAATGAGAGGAGGGTTTCTCCACTGATGCTGTCTGGATCTCTTCTGGCTTCAG 60
QY 108 TTCTGCAACGCGATGCCAAGTCACTTACAGAGAAGAAAACAGAGAACCCTTCGCCCC 167
DB 61 TTCTGCAACGCGAGGCCAAGTCACC-----TTACCGGAAAACAGAGAACCCTTCGCCCC 114
QY 168 AGAGGTGCTCCAGAGTTGTCAACGGAACCCGATGACTTTGAAGCAAAAAGGCATGCGAGT 227
DB 115 AGAGGTGCTCCAGAGTTGTCAACGGAACCCGATGACTTTGAAGCAAAAAGGCATGCGAGT 174
QY 228 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTG 287
DB 175 CTCGCTGCACCAAGCTCGAGTATGATCTCTGTGTCTATG-----ACACTG 222
QY 288 GCACCAACCAACCGTTCCCTCCAGGGAGCGGACACGTGGCCGCCCAACCCGAGACT 347
DB 223 GCGCCACCAACCAACCGTCCACCTCCAGGGAGCGGACACGTGGCCGCCCAACCCGAGACT 282
QY 348 ACGATGATGACCGCGTCAACCCGGAAGAGAGGAGCGCGATGGGACACAGCTGGAC 407
DB 283 ACGATGATGACCGCGTCAACCCGGAAGAGAGGAGCGCGATGGGACACAGCTGAAAC 342
QY 408 CGAGGAGCGGTGAAAGAGAGAAGACTGGAGACAACCAAGAGAGAAGATTGGAGCGACCA 467
DB 343 CGAGGAGCGGTGAAAGAGAGAAGACTGGAGACAACCAAGAGAGAAGATTGGAGCGACCA 402
QY 468 GTCATCAGCAGCCACGGAATAAAGCCCGAAGAGAGAGAGAGAGAACAGAGTGGGAA 527
DB 403 GTCATCAGCAGCCACGGAATAAAGCCCGAAGAGAGAGAGAGAACAGAGTGGGAA 462
QY 528 CACCGTAGCCATGTGAGGAGAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAA 587
DB 463 CACCGTAGCCAGGTGAGGAGAGAAACATCAGGAAACACCTTCTACTTCCCGTCAA 522
QY 588 GCGGTTTAGCACCCCGCTACCGGAAACCAAAACGGTAGGATCCGGGTCTCTGAGAGTTTG 647
DB 523 GCGGTTTAGCACCCCGCTACCGGAAACCAAAACGGTAGGATCCGGGTCTCTGAGAGTTTG 582
QY 648 ACCAAGGTCAAGGAGTTTCAGAAATCTCCAGATCACCGTATTTGTCAGATCGAGCCA 707
DB 583 ACCAAGGTCAAGGAGTTTCAGAAATCTCCAGATCACCGTATTTGTCAGATCGAGCCA 642
QY 708 AACCTAACACTCTTGTCTTCCCAAGCAGCTGATCTGATTAACATCTTGTATCCAGC 767
DB 643 GACCTAACACTCTTGTCTTCCCAAGCAGCTGATCTGATTAACATCTTGTATCCAGC 702
QY 768 AAGGCAAGCCACCGTAGCCAGTAAATGGAATTAACAGAAAGAGCTTTAATCTTGACG 827
DB 703 AAGGCAAGCCACCGTAGCCAGTAAATGGAATTAACAGAAAGAGCTTTAATCTTGACG 762
QY 828 AGGGCCATGCACTCAGAAATCCGTTTCAATTCCTACATCTTTGAACCGCATGACA 887
DB 763 AGGGCCATGCACTCAGAAATCCGTTTCAATTCCTACATCTTTGAATTCGACATGACA 822
QY 888 ACCAGAACCTCAGAGTAGCTAAATCTCCTGCGGTTAAACACACCGCGCAGTTTGAGG 947
DB 823 ACCAGAACCTCAGAGTAGCTAAATCTCCTGCGGTTAAACACCGCGCAGTTTGAGG 882
QY 948 ATTCTTCCCGGCGAGCAGCGAGACCAATCATCTTCTAGAGGCTTCAGAGGAATA 1007
DB 883 ATTCTTCCCGGCGAGCAGCGAGACCAATCATCTTCTAGAGGATTCAGAGGAATA 942
QY 1008 CGTTGAGGCGCGCTTCAATGCGGAATTCATGAGATACGGAGGTGCTGTTAGAGAGA 1067
DB 943 CTTTGAGGCGCGCTTCAATGCGGAATTCATGAGATACGGAGGTGCTGTTAGAGAGA 1002

QY 1068 ATGAGGAGGTGAGCAAGAGGAGAGAGCGGCAGAGCGATGAGTACTCGGAGTAGTGAGA 1127
DB 1003 ATGAGGAGGAGAGCAAGAGGAGAGAGAGCGGCAGAGCGATGAGTACTCGGAGTAGTG--- 1059
QY 1128 ACAATGAAGAGGTGATGATCAAAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTTAAGCACG 1187
DB 1060 ATATGAAGAGGTGATGATCAAAAGTGTCAAAGGAGCAGCTTCAAGAACTTACTTAAGCACG 1119
QY 1188 CTAATCCGTCTCAAAGAAAGGCTCCGAGAGAGGAGATATATCAACCAACCAATCAACT 1247
DB 1120 CTAATCCGTCTCAAAGAAAGGCTCC---GAAGAGAGAGATATATCAACCAACCAATCAACT 1176
QY 1248 TGAGAGAGGAGCGAGCCGATCTTTCTAACAATTTTGGGAAGTATTTGAGGTGAAGCCAG 1307
DB 1177 TGAGAGATGGGAGCGAGCCGATCTTTCTAACAATTTTGGGAAGTATTTGAGGTGAAGCCAG 1236
QY 1308 ACAAGAAAGAAACCCAGCTTCAGGACCTGGACATGATGCTCACTCTGTGTAGAGATCAAAG 1367
DB 1237 ACAAGAAAGAAACCCAGCTTCAGGACCTGGACATGATGCTCACTCTGTGTAGAGATCAAAG 1296
QY 1368 AAGGAGCTTTGATGCTCCCACTCACTTCAACTCAAAGGCCATGTTATCTGCTGTCAACA 1427
DB 1297 AAGGAGCTTTGATGCTCCCACTCACTTCAACTCAAAGGCCATGTTATCTGCTGTCAACA 1356
QY 1428 AAGGAACTGGAAACCTTTGAACTCTGCTGTGTAAGAAAGAGCAACAAGAGGAGCGGC 1487
DB 1357 AAGGAACTGGAAACCTTTGAACTCTGCTGTGTAAGAAAGAGCAACAAGAGGAGCGGC 1416
QY 1488 GGGAA-----GAAGAGGAGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1538
DB 1417 GGGAAAGAGAGTGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1476
QY 1539 TGGTAGGTAACAAGCGAGGTGAAGGAGGCGATGTTTCAATCATGCGAGCAGCTCATC 1598
DB 1477 TGGTAGGTAACAAGCGAGGTGAAGGAGGCGATGTTTCAATCATGCGAGCAGCTCATC 1536
QY 1599 CAGTAGCCATCAAGCTTCCCTCCGAACTCCATCTGCTTGGCTTGGTATCAAGCCTGAAA 1658
DB 1537 CAGTAGCCATCAAGCTTCCCTCCGAACTCCATCTGCTTGGCTTGGTATCAAGCCTGAAA 1596
QY 1659 ACAACCAAGAGATCTTCTTCAGGTGATAGGACAATGTATAGACAGATAGAGAGC 1718
DB 1597 ACAACCAAGAGATCTTCTTCAGGTGATAGGACAATGTATAGACAGATAGAGAGC 1656
QY 1719 AAGCGAAGGATTTAGCAATCTCCCTGGTGGGTGAACCAAGTTGAGAAGCTCATCAAAAACC 1778
DB 1657 AAGCGAAGGATTTAGCAATCTCCCTGGTGGGTGAACCAAGTTGAGAAGCTCATCAAAAACC 1716
QY 1779 AGAAGGAATCTCACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCCGTCGTC 1838
DB 1717 AGAAGGAATCTCACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCCGTC----- 1765
QY 1839 CTGAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1898
DB 1766 -----GTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1815
QY 1899 CACTCTTTCAATTTGAAGCTTTTAACTGAGAAATGGAGGCAACTTGTATGATCGAT 1958
DB 1816 CACTCTTTCAATTTGAAGCTTTTAACTGAGAAATGGAGGCAACTTGTATGATCGAT 1875
QY 1959 AATAAGATCAGCTTTTGTACTCTACTATCAAAAACTTATCAATAAATAAAGGTTG 2018
DB 1876 AATAAGATCAGCTTTTGTATCTACTATCAAAAACTTATCAATAAATAAAGGTTG 1935
QY 2019 TGGCTGTTTCTCC 2032
DB 1936 TGGCTGTTTCTCC 1949

RESULT 7

US-09-106-872A-18

; Sequence 18, Application US/09106872A

; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106.872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-09-106-872A-18

Query Match 36.9%; Score 750; DB 4; Length 750;

Best Local Similarity 100.0%; Pred. No. 1.3e-218; Indels 0; Gaps 0;
Matches 750; Conservative 0; Mismatches 0

QY 687 GTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTTGTTCTTCCCAAGCAGCTGATGCTG 746
DB 1 GTATTGTGCAGATCGAGGCCAAACCTTAACACTCTTTGTTCTTCCCAAGCAGCTGATGCTG 60
QY 747 ATAACTCTTGTATCCAGCAAGGCAAGCCAGCCGTAGCGTAGCAAAATGGCAATAACA 806
DB 61 ATAACTCTTGTATCCAGCAAGGCAAGCCAGCCGTAGCGTAGCAAAATGGCAATAACA 120
QY 807 GAAAGAGCTTTAATCTTTGAGGAGGCCATGACATCTGAGATCCCATCCGTTTCATTTCCT 866
DB 121 GAAAGAGCTTTAATCTTTGAGGAGGCCATGACATCTGAGATCCCATCCGTTTCATTTCCT 180
QY 867 ACATCTTGAACCCGATGACCAACCAAGAACTCTCAGAGTAGCTAAATCTCCATGCCGTTA 926
DB 181 ACATCTTGAACCCGATGACCAACCAAGAACTCTCAGAGTAGCTAAATCTCCATGCCGTTA 240
QY 927 ACACACCGGCGAGTTTGAGGATTTCTTCCCGGAGCAGCCGAGACCAATCATCTACT 986
DB 241 ACACACCGGCGAGTTTGAGGATTTCTTCCCGGAGCAGCCGAGACCAATCATCTACT 300
QY 987 TGAGGAGCTTTCAGCAGGAATACGTTGGAGGCCCTTCAATGCGGAATTCATGAGATAC 1046
DB 301 TGAGGAGCTTTCAGCAGGAATACGTTGGAGGCCCTTCAATGCGGAATTCATGAGATAC 360
QY 1047 GGAGGAGTCTTGTAGAGAGAAATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1106
DB 361 GGAGGAGTCTTGTAGAGAGAAATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 1107 GGAGTACTCGGAGTAGTGAGAAACAATGAAGAGTAGTAGTCAAAGTGTCAAAGGAGCAG 1166
DB 421 GGAGTACTCGGAGTAGTGAGAAACAATGAAGAGTAGTAGTCAAAGTGTCAAAGGAGCAG 480
QY 1167 TTGAAGAACTTACTAGCAGCTAAATCCGTTCAAGAAAGGCTCCGAGAGAGGAGGAG 1226
DB 481 TTGAAGAACTTACTAGCAGCTAAATCCGTTCAAGAAAGGCTCCGAGAGAGGAGGAG 540
QY 1227 ATATCAACCAACCAATCAACTTGAAGAGAGGCGAGCCGATCTTTCTAAACATTTTGGGA 1286
DB 541 ATATCAACCAACCAATCAACTTGAAGAGAGGCGAGCCGATCTTTCTAAACATTTTGGGA 600
QY 1287 AGTTATTTGAGGTGAAGCCAGACAAGAAACCCCAAGCTTTCAGGACCTGGACATATGC 1346
DB 601 AGTTATTTGAGGTGAAGCCAGACAAGAAACCCCAAGCTTTCAGGACCTGGACATATGC 660
QY 1347 TCACCTGTGTAGAGATCAAGAGAGGAGCTTTGATGCTCCCACTTCAACTCAAGGCCA 1406

Db 661 TCACCTGTGTAGAGTCAAGAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCCA 720
 Qy 1407 TGGTTATCGTCGTCTCAACAAAGGAAGT 1436
 Db 721 TGGTTATCGTCGTCTCAACAAAGGAAGT 750
 RESULT 8
 US-09-108-010B-3
 ; Sequence 3, Application US/09108010B
 ; Patent No. 6362399
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; GARY MICHAEL FADER
 ; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 ; CLASSES OF SOYBEAN SEED
 ; PROTEIN GENES
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 ; STREET: 1007 MARKET STREET
 ; CITY: WILMINGTON
 ; STATE: DELAWARE
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 19898
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: DISKETTE, 3.50 INCH
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: MICROSOFT WINDOWS 95
 ; SOFTWARE: MICROSOFT WORD 7.0A
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/108,010B
 ; FILING DATE: 30-Jun-1998
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/019,940
 ; FILING DATE: JUNE 14, 1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: LYNN M. CHRISTENBURY
 ; REGISTRATION NUMBER: 30,971
 ; REFERENCE/DOCKET NUMBER: BB-1071-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 302-992-5481
 ; TELEFAX: 302-773-0164
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1320 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 US-09-108-010B-3
 Query Match 20.0%; Score 406.8; DB 3; Length 1320;
 Best Local Similarity 62.1%; Pred. No. 1.9e-113;
 Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
 Qy 562 GAACAAACCTTTTACTTT---CCGCTCAAGCGGTTTAGCACCGCTACGGGAACCAAAA 618
 Db 90 GAATAACCTTTTCTACTTTAGAGCTCTAACAGCTTCCAAACTCTCTTTGAGAACCACAA 149
 Qy 619 CGGTAGATCCGGTCTCGAGAGGTTTGACCAAGGCTCAAGGAGTTTCAGATCTCCA 678
 Db 150 CGTTCCGATTCGTCTCTCCAGAGATTCAACAAACGCTCCCACTTCCCAACCTTGG 209
 Qy 679 GAATCACCGATTCTCGATTTCAGTCAAAACCCACACAACTCTTCCCAAGCACGC 738
 Db 210 AGACTACCGGATTCTCGATTTCAGTCAAAACCCACACAACTCTTCCCAAGCACGC 269
 Qy 739 TGATGCTGATAAATCTCTTTGTTATCCAGCAAGGCAAGCCGCTGACCGTAGCAATGG 798
 Db 270 TGACCGCGATTTCCTCTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 329

RESULT 9
 US-09-758-652-3
 ; Sequence 3, Application US/09758652

Qy 799 CAATAACAGAAAGAGCTTTAAATCTTGACGAGGGCCATGCACTCAGAAATCCCATCCGGTTT 858
 Db 330 CGACGACAGAGACTCTCTACAACTTCACTGCGGATGCCAGAGAAATCCAGCTGGAAC 389
 Qy 859 CATTTCCTCATCTTGACCGCCATGACAAACCAAGAACTCAGAGTAGCTAAATATCTCCAT 918
 Db 390 CACTTACTATTGTTTAAACCTCAGACCAACAGAAATCTCAAAATATCAAACTTGCCAT 449
 Qy 919 GCCCGTTAACACACCCCGCCAGTTTGAGGATTTCTTCCCGGCGAGCAGCGAGACCAATC 978
 Db 450 ACCCGTCAACAAACCTGCGAGATATGATGATTTCTTCTTATCTAGCACTCAAGCCCAACA 509
 Qy 979 ATCTTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGGCGCCTTCAATCGGAAATTCAA 1038
 Db 510 GTCTTACTTGCAGGGCTTCAGCAGGAATATCTAGAGACCTCTTCTCATAGCGAATTCGA 569
 Qy 1039 TGAGATACCGAGGGTCTGTTAAGAGAAATGACGAGGTGACAAAGAGAGAGAGAGGCA 1098
 Db 570 GGAGATAAACAGGGTTTGTGAGAGGAAGAGGAGCAGAGCGCAGCAGAG----- 621
 Qy 1099 GAGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTATAGTCAAGTGTCAAA 1158
 Db 622 -----GGAGTGTGATCGTGGAACTCTCAAA 644
 Qy 1159 GGAGCAGCTTGAAGAACTTACTAAGCACGCTAAATCCGCTCTCAAGAAAGGCTCCGAAGA 1218
 Db 645 GGACAAATTCGCAACTGAGCAGAGCTGCCAATCTAGTTCAAGGAAACCAATTTCTCTC 704
 Qy 1219 AGAGGAGATATCACCAACCCCAATCAACTTGAGAGAGGCGAGCCCGATCTTTCTAACAA 1278
 Db 705 CGAAGATG-----AACCAATTCAACTTGAGAGCGCAACCCCATCTTATTTCCAACAA 755
 Qy 1279 CTTTGGAGATTTATTTGAGTGAAGCCAGACAAAGAAAGAAACCCCGCTTCAGGACCTGGA 1338
 Db 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAACCCACAGCTTCGGAGCTGGA 812
 Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGAGCTTTGATGCTGCCACACTTCAACTC 1398
 Db 813 TATCTCTCTCAGTTCTGTGATATCAACGAGAGCTCTTCTTACCACACTTCAATTC 872
 Qy 1399 AAAGGCCATGTTATCGTCTCAACAAAGAACTTGAAACCTTGAACCTCTGGCTGT 1458
 Db 873 AAAGGCCATGATGATAGTATTAATGAAGGAGATGCAAAACATTTGAACTTGTGGCAT 932
 Qy 1459 ACAAAGAGACACACAGAGGGGACGGCGGGAAGAGAGGAGGAGGACGAAGACCGAAGAGA 1518
 Db 933 TAAAGAAACAAACAGAGCAGAAACAGGAAGAGGAACTTTG----- 975
 Qy 1519 GGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAAGCGATGTGT 1578
 Db 976 -----GAGTGAAGGTACAGAGCTGAATGCTCTGAAGACGATGATT 1019
 Qy 1579 CATCATGCCAGCAGCTCATCCAGTAGCCATCAACGTTCTCCGAACTCCATCTGTTGG 1638
 Db 1020 TGTAAATTCAGCAGCTTATCCATTTGCTGCTCAACGCTACCTCAAACTCAATTTCTTGC 1079
 Qy 1639 CTTGCGTATCAACGCTGAAACCAACACAGAAATCTTCTTGCAGGTGATAAGGACAATGT 1698
 Db 1080 TTTTGGTATCAATGCTGAGAACCAACAGAGGAACTTCTTCTGAGGCGAGAAACAAATGT 1139
 Qy 1699 GATAGACCATAGATAGAGAAACAGGAGGATTTAGCAATTCCTGGGTGGGTGAAACAAAT 1758
 Db 1140 GGTAAAGCAGATAGAAAGCAAGTGCAGGAGCTTGGTTCCCTGGGTCTGCACAAGATGT 1199
 Qy 1759 TGAGAGCTCATCAAAACACAGAAAGAAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
 Db 1200 TGAGAGGCTATTAAAGAGCAGAGGGAATCCTACTTTGTTGATGCTCAGCCTCA 1253

Patent No. 6703544
 GENERAL INFORMATION:
 APPLICANT: ANTHONY JOHN KINNEY
 GARY MICHAEL FADER
 TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 CLASSES OF SOYBEAN SEED
 PROTEIN GENES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 STREET: 1007 MARKET STREET
 CITY: WILMINGTON
 STATE: DELAWARE
 COUNTRY: UNITED STATES OF AMERICA
 ZIP: 19898
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.50 INCH
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: MICROSOFT WINDOWS 95
 SOFTWARE: MICROSOFT WORD 7.0A
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/758,652
 FILING DATE: 11-Jan-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/019,940
 FILING DATE: JUNE 14, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: LYNN M. CHRISTENBURY
 REGISTRATION NUMBER: 30,971
 REFERENCE/DOCKET NUMBER: BB-1071-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-5481
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1320 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 20.0%; Score 406.8; DB 4; Length 1320;
 Best Local Similarity 62.1%; Pred. No. 1.9e-113;
 Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
 562 GAACACCCCTTTCTACTT---CCGCTCAAGCGGTTTACAGCTTCAAACTCTCTTTGAGAACCAAA 618
 90 GAATAACCCCTTTCTACTTTAGAGCTTAAAGCTTCAAGCTTCCAACTCTCTTTGAGAACCAAA 149
 619 CGGTAGGATCCGGTCTCTCGAGAGGTTTGACAAAGGTCAAGGAGTTTCAAGATCTCCA 678
 150 CGTTCCGATTGCTCTCTCCAGAGATTCAACAAAGCTCCCACTTGAACCTTGAACCTTCG 209
 679 GAATCAGGATTTGAGAGATCGAGGCAACCTTAACTCTTCTTCTTCCCAAGCAGC 738
 210 AGACTACCGGATTTGCTGAGTTTCACTCAAAACCAACCAATCTCTTCCCCCAACATGC 269
 739 TGATGCTGATAACATCTTGTATTCAGCAGGCGGACCGCTGACCGTAGCAATGG 798
 270 TGACCGGATTTCTCTCTTTTGTCTTGTAGCGGAGAGCCACTTACCTTGTGTGAACAA 329
 799 CAATAACAGAAAGAGCTTTAATCTTGACGGGGCATGCACTCAGAACTCCCATCCGGTTT 858
 330 CGACGACAGAGACTCTTACAACTTCACTTGGGATGCGCAGAGATCCAGCTGGAAC 389
 859 CATTCTCTACATCTTGAAACCGCCATGACCAACAGAACTCAGATAGCTAAATCTCCAT 918
 390 CACTTACTATTGTTGTTAACTCTACGACCAACCAAGATCTCAAAATAATCAAACTTGCAT 449
 919 GCCCGTTAAACACACCCCGGCGAGTTTGTAGGATTTCTTCCCGGAGAGCGGAGACCAATC 978

Db 450 ACCGCTCAACAACTCGCAGATATGATGATTTCTTCTATCTAGCATCTAAGCCCAACA 509
 Qy 979 ATCTACTTTCAGGGCTTTCAGCAGGAATACGTTTGGAGCGCGCTTCAATCGGAATTCAA 1038
 Db 510 GTCCTACTTTCAGGGCTTTCAGCAGGAATATATCTAGAGACCTCTCTTCCATAGCAATCGA 569
 Qy 1039 TGAGATACGGAGGCTGCTTTAGAGAGAAATGCAAGGAGGTGAGCAAGAGGAGAGGGCA 1098
 Db 570 GGAGATAAACAGGGTTTGTGTTGGAGAGGAAGAGGAGCAGAGCAGCAAGAG----- 621
 Qy 1099 GAGCGATGAGTACTCGGAGTACTGAGAAACAATGAAGGAGTATAGTCAAAAGTGTCAAA 1158
 Db 622 -----GGAGTGATCTGGACTCTCAAA 644
 Qy 1159 GGAGCAGCTTGAAGAACTTAAAGCAGCTAAATCCGTCTCAAGAAAGGCTCCGAAGA 1218
 Db 645 GGAACAAATTCGGCAACTGAGCAGACGTCGCAAAATCTAGTTCAAGGAAACCAATTTCTCTC 704
 Qy 1219 AGAGGAGATATCAACCAACCAATCACTTGAAGAGGCGGAGCCGATCTTTCTTAACAA 1278
 Db 705 CGAAGATG-----ARCCATTTCAACTTTGAGAAGCGCCAAACCCCATCTATTCCCAACA 755
 Qy 1279 CTTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAAAGCAACCCCGAGCTTCAGGACCTGGA 1338
 Db 756 CTTTGGGAAGTTCTTTGAGATCACCCCTG---AGAAAAACCAAGCTTCGGGACTTGA 812
 Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGTCTCCACACTTCAACTC 1398
 Db 813 TATCTCTCTCAGTTCTGTGGATATCAACGAAGGAGCTCTTCTTCTACCACTTCAATTC 872
 Qy 1399 AAAGGCCATGTTATCGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1458
 Db 873 AAAGGCCATGTCATCTAGTCTGATTTAATGAAGGAGATGCAAAACATTTGAATTTGGCAT 932
 Qy 1459 AAGAAAGAGCAACAACAGAGGCGGCGGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1518
 Db 933 TAAAGAAACAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 975
 Qy 1519 GGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTTGAAGGAGGAGGAGGAGG 1578
 Db 976 -----GAAGTCAAGGATACAGAGCTGAAATTTGCTGAAGACGATGATTT 1019
 Qy 1579 CATCATCGCAGCAGCTCATCCAGTAGCCATCAAGCTTCTCCGAACTCCATCTGCTTGG 1638
 Db 1020 TGAATTTCCAGCAGCTTATCCATTTGTCTCAACGCTACCTCAAACTCAATTTCTTTC 1079
 Qy 1639 CTTTCGGTATCAACGCTGAAACCAACACAGAACTCTTCTTGCAGGTGATAAGGACAATGT 1698
 Db 1080 TTTTGGTATCAATGCTGAGAACCAACAGAGAACTCTTCTTGCAGGCGGAGGAGGAGGAG 1139
 Qy 1699 GATAGACCGATAGAGAACGAGGAGGATTTAGCATTTCCCTGGGTGCGGTGAACAGT 1758
 Db 1140 GGTAAAGCAGATAGAAAGCAAGTGCAGGAGCTTGGTTCCTTGGGTCTGCACAGATGT 1199
 Qy 1759 TGAGAGCTCATCAAAAACAGAGGAAATCTCACTTTGTGAGTGTCTGCTCTCA 1812
 Db 1200 TGAGAGGCTATTAAAGAGCAGAGGGAATCTTCTTGTGATGCTCAGCCTCA 1253

RESULT 10
 US-10-684-651-3
 ; Sequence 3, Application US/10684651
 ; Patent No. 6828491
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; GARY MICHAEL FADER
 ; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 ; CLASSES OF SOYBEAN SEED
 ; PROTEIN GENES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

OPERATING SYSTEM: MICROSOFT WINDOWS 95

SOFTWARE: MICROSOFT WORD 7.0A

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/108,010B

FILING DATE: 30-Jun-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/019,940

FILING DATE: JUNE 14,1996

ATTORNEY/AGENT INFORMATION:

NAME: LYNNE M. CHRISTENBURY

REGISTRATION NUMBER: 30,971

REFERENCE/DOCKET NUMBER: BB-1071-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: 302-992-5481

TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1818 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-108-010B-1

Query Match 19.9%; Score 404; DB 3; Length 1818;

Best Local Similarity 61.2%; Pred. No. 1.7e-112;

Matches 765; Conservative 0; Mismatches 410; Indels 75; Gaps 4;

```
Qy 563 AACAAACCTTTCTACTCCCGTCGAAGCGGTTTAGCACCCGCTAGCGGAAACCAACGGT 622
Db 577 AAGAAACCTTTCTCTCGGCTCTAACAGGTTTGAACCTCTCTTCAAAAACCAATATGGT 636
Qy 623 AGGATCCGGGCTCTCAGAGGTTTGACCAAAAGGTTCAAGGCGAGTTTTCAGAACTCTCCAGAA 682
Db 637 CGCATTCGGCTCTCCAGAGGTTCAACCAACGCTCCCAACAACTTCAGAACTCTCGAGAC 696
Qy 683 CACGATATGTCAGATCGAGGCGCAACCTAACACTCTTGTCTTCTTCCCAAGCAGCTGAT 742
Db 697 TACCGCATTTTGGAGTTTCAACTCCAAACCAACACCTCTCTCTCCCAACACTGCTGAC 756
Qy 743 GCTGATACATCTTGTATATCAGCAAGGCGCAACGCGTACCGTACCAATGGCAAT 802
Db 757 GCTGATACCTCATCGTTATCTTAAACGGAGCTGCCATTTCTTCTGTTGTAACCAACGAC 816
Qy 803 AACAGAAAGAGCTTTAATCTTGAAGGCGCATGCACTCAGAAATCCCATCCGTTTCAAT 862
Db 817 GACAGAGACTCTACAGACTTCAATCTGTGTATGCCCTGAGAGTCCCTCAGGAACACA 876
Qy 863 TCCTATACCTTTGAACCGCCATGACAAACAGAACTCAGAGTAGTCTTAAATCTCATGCCC 922
Db 877 TACTATGTTGGTCAACCCCTGACAAACAAAGAAATCTCAGATTAATAACAACTCGCCATACC 936
Qy 923 GTTAACACACCGCCAGTTTGAAGATTTCTTCCGGGAGCAGCGGAGCAATCATCC 982
Db 937 GTTAACAGCCTGGTAGATTGAGAGTTTCTTCTATCTAGCATGAGCTCAACAAATCC 996
Qy 983 TACTTGCAGGCTTCAGCAGGAATACGTTGAGGCGGCTTCAATGCGGAATTCAAATGAG 1042
Db 997 TACTTGCAGGATTCAGCAGGACATTTTGAAGGCTCTAGGATACCAATTCAGGAG 1056
Qy 1043 ATACGAGGGTGTGTTTGAAGAAGATGACAGAGGTGAGCAAGAGGAGAGGAGGAGG 1102
Db 1057 ATAAACAGGTTCTGTTTGTAGAGAGGAGGAGGAGGAGGAGGAGGAGGCTG--- 1113
Qy 1103 CGATGGAGTACTCGGAGTAGTGAGNACAAATGAAGAGTGATAGTCAAGTGTCAAGGAG 1162
Db 1114 -----CAAGAGAGCGGTGATTTGGAAATCTCGAAGAA 1146
Qy 1163 CACGTTGAAGAACTTACTAAGCAGCTAAATCCGTTCTCAAAAGAGGCTCCGAAGAAGAG 1222
Db 1147 CAGATTCCGGCACTGACGCAACAGTGCACCAATCTAGTTCAGAGAAACCAATTTCTTGAA 1206
```

```
Qy 1223 GGAGATATACCAACCCAAATCAACTTGAGAGAGGCGGAGCCGATCTTTCTAAACAACTTT 1282
Db 1207 G-----ATAAACCTTTTAACTTGAGAGGCGGAGCCCACTACTCTCAACAAGCTT 1257
Qy 1283 GGAAGTTTATTTGAGGTGAAGCAGACAAAGAAAGACCCCAAGCTTCAGGACCTCGACATG 1342
Db 1258 GGAAGTTTATTTGAGATCAACCCAG---AGAAAACCCCAAGCTTCGGGATTTGGATATC 1314
Qy 1343 ATGCTACCTGTGTAGAGATCAAAAGAGGAGCTTTGATGTCTCCACACTTCAACTCAAAAG 1402
Db 1315 TTCCTCAGTATTGTGGATATGAACGAGGAGCTCTCTCTTCTACCACTTCAATTTCAAAAG 1374
Qy 1403 GCCATGTTATCTGTCGTCGTCACAAAGAAAGAACTTGAACCTTGAACCTGCTGCTGAAGA 1462
Db 1375 GCGATAGTGATCTGGTAAATTAATGAAGGAGATGCAAACTTGAACCTTGTGGCTTAAAA 1434
Qy 1463 AAAGAGCAACAAACAGAGGCGGAGCGGAAAGAGGAGGAGGAGCAAGCAAGAGAGGAG 1522
Db 1435 GAACAAACACAGGAGCAGCAACAGGAAGAGCAACTTTG----- 1473
Qy 1523 GGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGAGGCGATGTGTTTCATC 1582
Db 1474 -----GAAGTGCAGAAATATAGAGCCGAATTGCTCTGAACAAGATATATTTGTA 1521
Qy 1583 ATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTCTGCTGGCTTC 1642
Db 1522 ATCCAGCAGGTTATCCAGTTGTGTCACGCTACCTCAAACTCTGAATTTCTTTGCTATT 1581
Qy 1643 GGTATCAACGCTGAAACAAACACAGAAATCTTCTTCGAGGTGATAAGGACAAATGTGATA 1702
Db 1582 GGTATTAATCCGAGAAACACAGAGGAACTTCTTCGAGGTTCCGAGCAAGCTGTTGAG 1641
Qy 1703 GACAGATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGTCGGGTGAACAAGTTGAG 1762
Db 1642 AGCCAGATACCTAGTCAAGTGCAGGAGTTGCAATTCCTCGGCTGCAAGCTGTTGAG 1701
Qy 1763 AGCTCATCAAAACCAAGAGGAATCTCATTGTTGAGTGTCTGCTCTCA 1812
Db 1702 AAGCTATTAAAGAACCAAGAGGAATCTTCTTTGTGATGCTCAGCCTAA 1751
```

RESULT 12

US-09-758-652-1

; Sequence 1, Application US/09758652

; Patent No. 6703544

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; APPLICANT: GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MICROSOFT WINDOWS 95

; SOFTWARE: MICROSOFT WORD 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/758,652

; FILING DATE: 11-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/019,940

; FILING DATE: JUNE 14,1996

; ATTORNEY/AGENT INFORMATION:

[illegible]

; LENGTH: 1818 base pairs	
; TYPE: nucleic acid	
; STRANDEDNESS: single	
; TOPOLOGY: linear	
; MOLECULE TYPE: cdna	
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
US-10-684-651-1	
Query Match 19.9%; Score 404; DB 4; Length 1818;	
Best Local Similarity 61.2%; Pred. No. 1.7e-112;	
Matches 765; Conservative 0; Mismatches 410; Indels 75; Gaps 4;	
Qy	563 AACAAACCTCTTCTACTTCCCGTCGAGGCGGTTTAGCACCCGCTAGCGGAAACCAAAACGGT 622
Db	577 AAGAAACCTCTTCTCTCGGCTCTAACAGGTTTGGAACTCTCTTCAAAAACCAATATGGT 636
Qy	623 AGGATCCGGGCTCTGACAGAGGTTTGACCAAAAGGTCAGGCGAGTTTTCAGAACTCTCCAGAAAT 682
Db	637 CGCATTCGCGTCTCCACAGAGTTCAACCAACGCTCCGCCACAACCTTCAGAAATCTCCGAGAC 696
Qy	683 CACGATATTGTCAGATCGAGCCAAACCTTAACACTCTTGTCTTCTTCCCAAGCAGCGCTGAT 742
Db	697 TACCGCATTTTGGAGTTCAACTCCAAACCCCAACACCCCTCTCTCTCCCAACCATGCTGAC 756
Qy	743 GCTGATAACATCTTGTATTCCAGCAAGGCAAGCCACCGTGACGCTAGCAAAATGGCAAT 802
Db	757 GCTGATTAACCTATCGTTATCTTAAACGGGACTGCTCATTTCTTCTTGGTGAACAACGAC 816
Qy	803 AACAGAAAGAGCTTTAATCTTTGACAGGCGGCATGCACTCAGAAATCCCATCCGTTTTCATT 862
Db	817 GACAGAGACTCTACAGACTTCAATCTGTCATGCTGAGTCCCTGAGAGTCCCTCAGGAACACA 876
Qy	863 TCCTACATCTTTGACCGCCATGACAAACAGAACTCTCAGAGTAGTGTAAATCTCAATGCC 922
Db	877 TACTATGTGTCAACCTCTGACAAACGAAATCTCAGATTAATAACCACTCGCCATACCC 936
Qy	923 GTTAACACACCGCCAGTTTGGAGTTTCTTCCGGCGAGCAGCGAGCAATCATCC 982
Db	937 GTTAACAGCTGTGTAGATTTGAGAGTTTCTTCTATCTAGCATGAAGCTCAACAATCC 996
Qy	983 TACTTGCAGGCTCTCAGCAGGAATACGTTGGAGCGCGCTTCAATGCCGAATTCAAATGAG 1042
Db	997 TACTTGCAGGATTCAGCAGGAACATTTTAGAGGCTCTTAGATACCAATTCGAGGAG 1056
Qy	1043 ATACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGAGCAAGAGGAGAGAGGAGGAG 1102
Db	1057 ATAAACAAGGTTCTGTTTAGTAGAGAGAGGCGAGCAGCAAGGCGGAGCAGAGCTG--- 1113
Qy	1103 CGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGGTAGTCAAGTCTCAAGGAG 1162
Db	1114 -----CAAGAGAGCGGTGATTGTGGAAATCTCGAAGGAA 1146
Qy	1163 CACGTTGAAGAACTTACTAAGCAGCTAAATCCGCTCTCAAGAAAGGCTCCGAAGAAGAG 1222
Db	1147 CAGATTGGGCACTCAGCAGAACTGCTGCTAGTTCAGGAAACCAATTTCTTCTGAA 1206
Qy	1223 GGAGATATCAACAACCCAAATCACTTTAGAGAGAGCGGAGCCCGATCTTTCTAAACACTTT 1282
Db	1207 G-----ATAAACCTTTTAACTTGAGAGCGCGCACCCCATCTACTCCCAACAAGCTT 1257
Qy	1283 GGAAGTTATTGTAGGTGAAGCCAGACAGAAAGAACCCCGAGCTTCAGGACCTCGACATG 1342
Db	1258 GCGAAGTCTTTAGATCACCCAG---AGAAAAACCCCGAGCTTCGGGACTTGGATATC 1314
Qy	1343 ATGCTCACTCTGTAGAGATCAAGAAGGAGCTTTGATGCTCCCACTTCAACTCAAAAG 1402
Db	1315 TTCTCTCAGTATTGTGATATGACGAGGAGCTCTTCTTCTTACCACTTCAATTCAAAG 1374
Qy	1403 GCCATGGTTATGTCGTCGTCAACAAGAACTTGAACCTTGAACCTGTCGCTGTGAAGA 1462
Db	1375 GCGATAGTGA TACTGGTAAATTAATGAGGAGATGCAACACTTGAACCTTGTGGCTTAAA 1434
Qy	1463 AAAGAGCAACACAGAGGGGAGCGGGGGAAGAGAGGAGGAGGAGCAAGAGAGAGAG 1522
Db	1435 GAACAACAACGAGGAGCAGCAACAGGAAGAGCAACCTTTG----- 1473
Qy	1523 GGAAGTAACAGAGAGGTCGCTAGGTACACAGCGAGGTTTGAAGAGGCGATGTGTTCATC 1582
Db	1474 -----GAAGTGGGAAATATAGAGCCGAATTTGTCTGAACAAGATATATTTGTA 1521
Qy	1583 ATGCGACAGCTCATCCAGTAGCCATCAACGCTTCTCCGNACTCCATCTGCTTGGCTTC 1642
Db	1522 ATCCAGCAGGTTATCCAGTTGTGTCAACGCTACCTCAAACTCTGAATTTCTTTGCTATT 1581
Qy	1643 GGATCAACGCTGAAAAACAACACAGAAATCTTCTTGCAGGTGATAGGACAAATGTGATA 1702
Db	1582 GGTATTATGCGGAGAACACACAGAGGAACTTCTCTCGCAGGTTCCGACAGCAATGTGATA 1641
Qy	1703 GACCAGATAGAGAAGCAAGCAAGGATTTAGCATTCCTCGGTCGGGTGAAACAAGTTGAG 1762
Db	1642 AGCCAGATACCTAGTCAAGTGCAGGAGTTTGCAATTCCTCGCAACAGCTGTTGAG 1701
Qy	1763 AGCTCATCAAAAACCAACAGAGGAATCTCACTTTTGTGAGTCTGCTCTCTCA 1812
Db	1702 AAGCTATTAAAGAACCAAGAGAAATCTTACTTTGTGATGCTCAGCCTAA 1751
RESULT 14	
US-09-108-010B-2	
; Sequence 2, Application US/09108010B	
; Patent No. 6362399	
; GENERAL INFORMATION:	
; APPLICANT: ANTHONY JOHN KINNEY	
; GARY MICHAEL FADER	
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC	
; CLASSES OF SOYBEAN SEED	
; PROTEIN GENES	
; NUMBER OF SEQUENCES: 22	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY	
; STREET: 1007 MARKET STREET	
; CITY: WILMINGTON	
; STATE: DELAWARE	
; COUNTRY: UNITED STATES OF AMERICA	
; ZIP: 19898	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: DISKETTE, 3.50 INCH	
; COMPUTER: IBM PC COMPATIBLE	
; OPERATING SYSTEM: MICROSOFT WINDOWS 95	
; SOFTWARE: MICROSOFT WORD 7.0A	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: US/09/108,010B	
; FILING DATE: 30-Jun-1998	
; CLASSIFICATION: <Unknown>	
; PRIOR APPLICATION DATA:	
; APPLICATION NUMBER: 60/019,940	
; FILING DATE: JUNE 14, 1996	
; ATTORNEY/AGENT INFORMATION:	
; NAME: LYNNE M. CHRISTENBURY	
; REGISTRATION NUMBER: 30,971	
; REFERENCE/DOCKET NUMBER: BB-1071-A	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: 302-992-5481	
; TELEFAX: 302-773-0164	
; INFORMATION FOR SEQ ID NO: 2:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 1920 base pairs	
; TYPE: nucleic acid	
; STRANDEDNESS: single	
; TOPOLOGY: linear	
; MOLECULE TYPE: cdna	
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
US-09-108-010B-2	
Query Match 16.0%; Score 324.2; DB 3; Length 1920;	
Best Local Similarity 59.5%; Pred. No. 4.7e-86;	

Matches 764; Conservative 0; Mismatches 408; Indels 135; Gaps 6;	
Qy	563 AACAAACCTTTCTACTCCGCTCAAGCGGTTTAGCACCGCTACCGGAAACCAAAACGGT 622
Db	625 AAGAACCTTTTCACTTCAACTCTAAAGGTTCCAAACTCTTCAAAAACCAATATGGC 684
Qy	623 AGAATCCGGGCTCTGACAGGTTTGACAAAGGTCGAAGGCAAGTTTCAGATCTCCAGAA 682
Db	685 CACGTTCCGGCTCTCCAGAGGTTCAACAAACGCTCCCAACAGCTTCAAGATCTCCGAGAC 744
Qy	683 CACGTTATGTCAGATCGAGGCCAAACCTTAACACTCTTGTCTTCCCAAGCAGCTGAT 742
Db	745 TACCGCATTTTGGAGTTCAACTCCAAACCCCAACACCTTCTTCTCCCCCAACATGCTGAC 804
Qy	743 GCTCATAACTCCTTTGTTATCCAGCAAGGGCAAGCCACCGTAGCGTAGCAAAATGGCAAT 802
Db	805 GCTGATTAACCTCATCGTTATCTTAACGGGACTGCCATTTCTTACCTTGGTGAACACGAC 864
Qy	803 AACAGAAAGAGCTTTAATCTTGAAGGCGCATGCACTCAGAAATCCC----- 849
Db	865 GACCGAGACTCTTACAACTTCAATCTGGCGATGCCCTAAGAGTCCCTGCAGGAACCA 924
Qy	850 -----ATCGGTTTCAATTCC 865
Db	925 TTCTATGTGGTTAAACCTGCAACACGACGAGAATCTCAGATGATGACGAGAAACACATTC 984
Qy	866 TACATCTTGAACCCCATGACCAACAGAACTCAGAGTAGCTAAATCTCCATGCCCGTT 925
Db	985 TATGTGTTAAACCTGACAGCAGAGAAATCTCAGATGATACACTCGGCATACCCGTT 1044
Qy	926 AACACAACCCGCGCTTTGAGGATTTCTTCCGCGCAGCAGCCGAGACCAATCATCTTAC 985
Db	1045 AACAAACCCGCTAGATTGAGAGTTTCTCTATCTAGCACTCAAGCTCAACAGTCCCTAC 1104
Qy	986 TTGCAGGCTTCAGCAGGAATACCTTGGAGGCCCTTCAATCGGAAATCAATGAGATA 1045
Db	1105 TTGCAAGGGTTTCAAGAAATATTTCTAGAGCCCTCATACACACCAAAATTCGAGGAGATA 1164
Qy	1046 CGGAGGCTGCTTTAGAAAGAAATGCAAGGAGGTGAGCAAGGAGGAGAGGCGAGAGCGGA 1105
Db	1165 AACAGGTTCTGTTGTTGAGAGGAGGGGAGCAACAAAGGGAGGAGAGCTGC----- 1219
Qy	1106 TGGAGTACTCGGAGTAGTGAGAA CAATGAAGAGTAGTAGTCAAAAGTGTCAAAGGAGCAC 1165
Db	1220 -----AAGAGAGTGTGATTGTTGGAATCTCAAGAAACAA 1254
Qy	1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGTTCAAGAAAGGCTCCGAGAGAGGGA 1225
Db	1255 ATTCCGGAACCTGACCAAAACATGCCAAATCTAGTTCAAGGAAACCATTTCTTCTGAAG-- 1312
Qy	1226 GATATCAACCAACCAATCAACTTCAGAGAGGCGGAGCCGATCTTTCTAAACACTTTGGG 1285
Db	1313 -----ATAACCTTTCAACTTGGGAAGCGCGAACCCCATCTATTCCAACAGCTTGGC 1365
Qy	1286 AAGTTATTGAGGTGAAGCCAGACAAGAAAGACCCCGCTTCAGGACTTGAAGCATGATG 1345
Db	1366 AAGTTGTTGAGATTACCC-----AGAGAAACCTCAGCTTCGGGACTTGGAGTCTTC 1419
Qy	1346 CTCACCTGTGTAGAGATCAAGAGAGGAGCTTTGATGCTCCACACTTCAACTCAAGGCC 1405
Db	1420 CTCAGTGTGTGGATATGAACAGGGGAGCTTTTCTTACACACTTCAATTCAAAGGCC 1479
Qy	1406 ATGTTTATCGTCGTCGTCAACAAAGGAACTGGAAACCTTGAACCTCGTGGCTGTGAAGAAA 1465
Db	1480 ATAGTGTACTAGTGTATTATGAAGAGAGAGCAACATTTGACTTGT---TGGCATTTAA 1536
Qy	1466 GAGCAACAAACAGAGGGAGCGCGGGAAGAGAGAGGAGCAAGAGCAAGAGAGAGGAGGA 1525
Db	1537 GAAACAAACACAGAGGCGAGCAACAGAGAGAGCAACCTTTGGAA----- 1578
Qy	1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGTAAGGAGCGGATGTGTTTCATATG 1585
Db	1579 -----GTGCGGAAATATAGAGCTGAATTGTCTGAACCAAGATATATTGTGAATC 1626

Qy	1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db	1627 CCAGCAGGTTATCCAGTTATGTCACGCTACCTCAGATCTGAATTTCTTTGCTTTGGT 1686
Qy	1646 ATCAACGCTGAAAAACCAACACAGAAATCTTCTTGCAGGTGATAGGACAAATGTGATAGAC 1705
Db	1687 ATCAATGCCGAGAACCAACGAGAGAACTTCTTTCAGAGTTTCGAAAGACAATGTGATAAGC 1746
Qy	1706 CAGATAGAGAAGCAAGCGAGGATTTAGCAATCCCTGGGTCGGTGAACAAGTTGAGAAG 1765
Db	1747 CAGATACCTAGTCAAGTGCAGGAGCTTGCCTTCCCTAGGCTCGAAAAAGATATTGAGAAC 1806
Qy	1766 CTCAATCAAAAAACAGAGGAATCTCACTTTGTGAGTCTCGTCTCTCA 1812
Db	1807 CTAAATAAGAGCCAAAGTGAGTCTACTTTGTGATGCTCAGCCTCA 1853
RESULT 15	
US-09-758-652-2	
; Sequence 2, Application US/09758652	
; Patent No. 6703544	
; GENERAL INFORMATION:	
APPLICANT: ANTHONY JOHN KINNEY	
GARY MICHAEL FADER	
TITLE OF INVENTION: SUPPRESSION OF SPECIFIC	
CLASSES OF SOYBEAN SEED	
PROTEIN GENES	
NUMBER OF SEQUENCES: 21	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY	
STREET: 1007 MARKET STREET	
CITY: WILMINGTON	
STATE: DELAWARE	
COUNTRY: UNITED STATES OF AMERICA	
ZIP: 19888	
COMPUTER READABLE FORM:	
MEDIUM TYPE: DISKETTE, 3.50 INCH	
COMPUTER: IBM PC COMPATIBLE	
OPERATING SYSTEM: MICROSOFT WINDOWS 95	
SOFTWARE: MICROSOFT WORD 7.0A	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/09758,652	
FILING DATE: 11-Jan-2001	
CLASSIFICATION: <Unknown>	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 60/019,940	
FILING DATE: JUNE 14,1996	
ATTORNEY/AGENT INFORMATION:	
NAME: LYNNE M. CHRISTENBURY	
REGISTRATION NUMBER: 30,971	
REFERENCE/DOCKET NUMBER: BB-1071-A	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: 302-992-5481	
TELEFAX: 302-773-0164	
INFORMATION FOR SEQ ID NO: 2:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 1920 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: cDNA	
SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
US-09-758-652-2	

Query Match	
Best Local Similarity 58.5%;	
Matches 764; Conservative 0; Mismatches 408; Indels 135; Gaps 6;	
Qy	563 AACAAACCTTTCTACTCCGCTCAAGCGGTTTAGCACCGCTTACGGGAAACCAAAACGGT 622
Db	625 AAGAACCTTTTCACTTCAACTCTAAAGGTTCCAAACTCTTCAAAAACCAATATGGC 684

Qy 623 AGGATCCGGTCTCGACAGGTTTGACCAAGGTCAAGGAGTTTCAGAAATCTCCAGAA 682
Db 685 CACGTTCCGTCCTCCAGAGTTTCAACAAAGCTCCCAACAGCTTCAGAAATCTCCGAGAC 744
Qy 683 CACGTAATGTGACAGATCGAGGCCAAACTTAACACTCTTTGTCTTCCCAAGCAAGCTGAT 742
Db 745 TACGCAATTTGGAGTTCAACTCCAAACCCCAACACCTTCTTCTCCCCCAACCATGCTGAC 804
Qy 743 GCTGATAACATCTTGTATTCAGCAAGGGCAAGCCAGCTGACGGTAGCAAAATGGCAAT 802
Db 805 GCTGATTAACCTCATCTGTTATCTTAACGGGACTGCCAATCTTACCTTTGGTGAACAACGAC 864
Qy 803 AACAGAAAGAGCTTTAATCTTTCAGAGGGCCATGCACTCAGAATCCC----- 849
Db 865 GACCGAGACTCTTACAACCTTTCAATCTGGCGATGCCCTAAGAGTCCCTGCAGGAACCA 924
Qy 850 -----ATCCGGTTTCATTTCC 865
Db 925 TTCTATGTGGTTAAACCTTGACAACGACGAGAATCTCAGAATGATAGCAGGAACCAATTC 984
Qy 866 TACATCTTTGAACCGCATGACAACAGAACTCAGAGTAGCTAAATCTCATGCCCCGTT 925
Db 985 TATGTGGTTAAACCTTGACAACGAGAGAACTCAGAATGATAACACTCGCCATACCCGTT 1044
Qy 926 AACACACCCGCCAGTTTGGAGTTTCTTCCCGGAGCGAGCCGAGACCAATCATCCTAC 985
Db 1045 AACAAACCCGTAATTTGAGAGTTTCTTCTATCTAGCACTCAAGCTCAACAGTCTCTAC 1104
Qy 986 TTGCAGGCTTCAGCAGGAATACGTTGGAGCCGCTTCAATGCGGAATCAATGAGATA 1045
Db 1105 TTGCAAGGGTTTCAGCAAGAATTTCTAGAGGCTCTATACGACACCAATTCGAGAGATA 1164
Qy 1046 CGGAGGTGCTGTTAGAAAGAAATGACAGAGGTGAGCAAGAGGAGAGAGGGCAGAGCGA 1105
Db 1165 AACAAAGGTCTGTTTGGTAGAGAGGGGCGAGCAACAAGGGGAGAGAGGCTGC----- 1219
Qy 1106 TGGAGTACTCGGAGTAGTGAGAACAAATGAAGAGTGATAGTCAAAAGTGTCAAAGGAGCAC 1165
Db 1220 -----AAGAGAGTGTGATTTGGAAATCTCAAGAAACAA 1254
Qy 1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAAGAAAGGCTCCGAAGAAAGAGGGA 1225
Db 1255 ATTCGGGAACCTGACAAACATGCCAAATCTAGTTCAAGGAAACCAATTTCTCTGAAG-- 1312
Qy 1226 GATATCACCAACCAATCACTTGAGAGAGGCGAGCCGATCTTTCTAACAACTTTGGG 1285
Db 1313 -----ATAAACCTTTCAACTTGGGAAGCGCGACCCCATCTATTCCAAACAGCTTGGC 1365
Qy 1286 AAGTTATTGAGGTGAAGCCAGACAGAAAGAACCCCGAGCTTCAGGACCTGGACATGATG 1345
Db 1366 AAGTTGTTTGAGATTACCC-----AGAGAAACCTTCAGCTTCGGGACTTGGATGTCTTC 1419
Qy 1346 CTCACCTGTGTAGAGATCAAGAAAGGAGCTTTGATGCTCCACACTTCAACTCAAAGGCC 1405
Db 1420 CTCAGTTGTGGATATGAACGAGGAGCTCTTTTCTACACACTTCAATTCAAAGGCC 1479
Qy 1406 ATGTTATPGTGTGTCACAAAGGAACCTGGAACCTTGAACTCGTGGCTGTAAAGAAA 1465
Db 1480 ATAGTGTACTAGTGATTAATGAAGGAGAACCAACATTGAACTTGT---TGGCATTA 1536
Qy 1466 GAGCAACACAGAGGGGACCGGGGAGAGAGGAGGAGCGAGAGAGAGAGAGAGGGA 1525
Db 1537 GAACAAACACAGAGGAGCAGCAACAGGAAGAGCACTTTGGAA----- 1578
Qy 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTTCATCATG 1585
Db 1579 -----GTGCGGAATATAGAGCTGAATTTGTCTGAACAGATATATTGTAAATC 1626
Qy 1586 CCAGGAGCTATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db 1627 CCAGCAGGTTATCCAGTTATGGTCAACGCTACCTCAGATCTGAATTTCTTTGCTTTGGT 1686
Qy 1646 ATCAACGCTGAACCAACCAAGAAATCTTCTTCCAGGTGATAGGACAATGTGATAGAC 1705

Db 1687 ATCAATGCCGAGAACCAACGAGGAACCTTCTTGACAGTTTCGAAAGACAATGTGATAAGC 1746
Qy 1706 CAGATAGAGAAAGCAAGCGAAGGATTTAGCAATTCCTGGGTGGGTGAACAAGTTGAGAAG 1765
Db 1747 CAGATACCTAGTCAAGTGCAGGAGCTTGCCTTAGTCTGCAAAAGATATTGAGAAC 1806
Qy 1766 CTCAATCAAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGTCCTCA 1812
Db 1807 CTAATAAAGAGCCAAAGTGAGTCTCTACTTTTGTGGATGCTCAGCCTCA 1853

Search completed: August 24, 2005, 03:56:27
Job time : 348.307 secs

This Page Blank (uspiu)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 1228.33 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 2032
Sequence: 1 aataatcatatatttcac.....cgtttgctgctgtttctcc 2032

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2032	100.0	2032	10	US-09-731-375A-1
2	2032	100.0	2032	14	US-10-228-806-1
3	2032	100.0	2032	17	US-10-100-303A-5
4	2032	100.0	2032	20	US-10-728-323-1
5	2032	100.0	2032	20	US-10-728-051-1
6	2032	100.0	2032	21	US-10-899-551-1
7	2032	100.0	2032	21	US-10-958-324-5

8	1752.4	86.2	1949	17	US-10-100-303A-6	Sequence 6, Appli
9	406.8	20.0	1254	17	US-10-245-227B-15	Sequence 15, Appli
10	406.8	20.0	1278	17	US-10-245-227B-13	Sequence 13, Appli
11	406.8	20.0	1320	9	US-09-758-652-3	Sequence 3, Appli
12	406.8	20.0	1320	17	US-10-245-227B-3	Sequence 3, Appli
13	406.8	20.0	1320	18	US-10-684-651-3	Sequence 3, Appli
14	406.8	20.0	1320	19	US-10-757-074-3	Sequence 3, Appli
15	406.8	20.0	1320	19	US-10-757-155-3	Sequence 3, Appli
16	406.8	20.0	1320	19	US-10-757-667-3	Sequence 3, Appli
17	404	19.9	1818	19	US-09-758-652-1	Sequence 1, Appli
18	404	19.9	1818	18	US-10-684-651-1	Sequence 1, Appli
19	404	19.9	1818	19	US-10-757-074-1	Sequence 1, Appli
20	404	19.9	1818	19	US-10-757-155-1	Sequence 1, Appli
21	404	19.9	1818	19	US-10-757-667-1	Sequence 1, Appli
22	404	19.9	2073	20	US-10-739-930-3296	Sequence 3296, Ap
23	404	19.9	2260	18	US-10-424-599-10353	Sequence 10353, A
24	404	19.9	5642	18	US-10-424-599-10364	Sequence 10364, A
25	391	19.2	5642	18	US-10-424-599-10364	Sequence 10364, A
26	390	19.2	1656	18	US-10-425-114-9571	Sequence 9571, Ap
27	374.2	18.4	2031	18	US-10-424-599-117263	Sequence 117263,
28	324.2	16.0	1920	9	US-09-758-652-2	Sequence 2, Appli
29	324.2	16.0	1920	18	US-10-684-651-2	Sequence 2, Appli
30	324.2	16.0	1920	19	US-10-757-074-2	Sequence 2, Appli
31	324.2	16.0	1920	19	US-10-757-155-2	Sequence 2, Appli
32	324.2	16.0	1920	19	US-10-757-667-2	Sequence 2, Appli
33	293.4	14.4	1107	18	US-10-425-114-8117	Sequence 8117, Ap
34	279.8	13.8	1088	18	US-10-425-114-7645	Sequence 7645, Ap
35	279.4	13.7	1083	18	US-10-425-114-14774	Sequence 14774, A
36	279.2	13.7	1079	18	US-10-425-114-8143	Sequence 8143, Ap
37	267	13.1	1037	18	US-10-425-114-14858	Sequence 14858, A
38	263.8	13.0	1034	18	US-10-425-114-14773	Sequence 14773, A
39	260.4	12.8	1024	18	US-10-425-114-7686	Sequence 7686, Ap
40	247.8	12.2	1011	18	US-10-425-114-8555	Sequence 8555, Ap
41	233	11.5	951	18	US-10-425-114-8579	Sequence 8579, Ap
42	214.8	10.6	863	18	US-10-425-114-7590	Sequence 7590, Ap
43	214.8	10.6	880	18	US-10-425-114-7667	Sequence 7667, Ap
44	214.8	10.6	934	18	US-10-425-114-6929	Sequence 6929, Ap
45	199.6	9.8	720	18	US-10-425-114-15167	Sequence 15167, A

ALIGNMENTS

RESULT 1
US-09-731-375A-1
; Sequence 1, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731.375A
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-731-375A-1

Query Match	100.0%;	Score 2032;	DB 10;	Length 2032;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2032;	Conservative	0;	Mismatches	0;	
		Indels	0;	Gaps	0;
Qy	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60		
Db	1	AATAATCATATATATTCATCAATCATCTATATAGTAGTAGCAGGAGCAATGAGAGGAG	60		
Qy	61	GGTTTCTCCACTGATGCTGTGCTAGGATCCTTGCTGCTTCAGTTCTGCAACGCA	120		

Db 61 GTTTCCTCCACTGATGCTGTGTAGGATCTTTGTCTCTGGCTTCAGTTTCTGCAACGCA 120
Qy 121 TGCCAAAGTCTACCTTTACAGAAAGAAAAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCA 180
Db 121 TGCCAAAGTCTACCTTTACAGAAAGAAAAACAGAGAACCCCTGCGCCACAGAGGTGCCTCCA 180
Qy 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAAGCAAAAGGCGATGCGAGTCTCGCTGCACCAA 240
Db 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAAGCAAAAGGCGATGCGAGTCTCGCTGCACCAA 240
Qy 241 GCTCGAGTATGATCTCTGTTGTCTATGATCTCTGAGGACACATGCGCACCAACCAACCA 300
Db 241 GCTCGAGTATGATCTCTGTTGTCTATGATCTCTGAGGACACATGCGCACCAACCAACCA 300
Qy 301 ACCTTCCCTCCAGGAGCGGACACCTGCGCGCCCAACCCGAGACTACGATGATGACCG 360
Db 301 ACCTTCCCTCCAGGAGCGGACACCTGCGCGCCCAACCCGAGACTACGATGATGACCG 360
Qy 361 CCGTCAACCCCGAAGAGAGGAGGCGCGATGCGGACCCAGCTGGACCCGAGGAGCGTGA 420
Db 361 CCGTCAACCCCGAAGAGAGGAGGCGCGATGCGGACCCAGCTGGACCCGAGGAGCGTGA 420
Qy 421 AAGAGAAAGACTGAGAGACACCAAGAGAAAGATTGAGGCGACCAAGTCAATCAGCAGCC 480
Db 421 AAGAGAAAGACTGAGAGACACCAAGAGAAAGATTGAGGCGACCAAGTCAATCAGCAGCC 480
Qy 481 ACGGAAATTAAGGCCCGAAGAGAGAGAGAACAGAGTGGGGAACACACAGTAGCCCA 540
Db 481 ACGGAAATTAAGGCCCGAAGAGAGAGAGAACAGAGTGGGGAACACACAGTAGCCCA 540
Qy 541 TGTGAGGGAAGAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Db 541 TGTGAGGGAAGAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Qy 601 CCGTACCGGAAACCAAAACCGTAGGATCCGCGTCTCGAGAGTTTACCAAGGTCAG 660
Db 601 CCGTACCGGAAACCAAAACCGTAGGATCCGCGTCTCGAGAGTTTACCAAGGTCAG 660
Qy 661 GCAGTTTCAGAACTCTCAGAAATCACCGTATTTGTCAGATCGAGGCGAAACCTAACACTCT 720
Db 661 GCAGTTTCAGAACTCTCAGAAATCACCGTATTTGTCAGATCGAGGCGAAACCTAACACTCT 720
Qy 721 TGTTCCTTCCCAAGCAGCTGATGCTGATAACATCTCTGTTATCCAGCAAGGCGAACCCAC 780
Db 721 TGTTCCTTCCCAAGCAGCTGATGCTGATAACATCTCTGTTATCCAGCAAGGCGAACCCAC 780
Qy 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCAC 840
Db 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCAC 840
Qy 841 CAGAAATCCCATCCGGTTTCAATTTCTTACATCTTGAACCGGCATGACAAACAGAACTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCAATTTCTTACATCTTGAACCGGCATGACAAACAGAACTCAG 900
Qy 901 AGTAGCTAAATCTCCATCCCGTTTAAACACACCCCGGCGAGTTTGAAGATTTCTTCCCGGC 960
Db 901 AGTAGCTAAATCTCCATCCCGTTTAAACACACCCCGGCGAGTTTGAAGATTTCTTCCCGGC 960
Qy 961 GAGCAGCGGAGACCAATCTACTTTCAGAGGCTTCAGCAGGAATACGTTGAGGCGGC 1020
Db 961 GAGCAGCGGAGACCAATCTACTTTCAGAGGCTTCAGCAGGAATACGTTGAGGCGGC 1020
Qy 1021 CTTCAATCCGGAATTTCAATGAGATACGAGGCTGCTTTAGAGAGAAATGACGAGGTGA 1080
Db 1021 CTTCAATCCGGAATTTCAATGAGATACGAGGCTGCTTTAGAGAGAAATGACGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGCGAGGCGATCGAGTACTCGAGTAGTGAGAAACAAATGAAGAGT 1140
Db 1081 GCAAGAGGAGAGGCGAGGCGATCGAGTACTCGAGTAGTGAGAAACAAATGAAGAGT 1140
Qy 1141 GATAGTCAAAAGTGTCAAGAGGACGTTGAGACTTACTTAAGCAGCTAAATCCGCTC 1200
Db 1141 GATAGTCAAAAGTGTCAAGAGGACGTTGAGACTTACTTAAGCAGCTAAATCCGCTC 1200

Qy 1201 AAAAGAAAGGCTCCGAAGAAAGAGGAGATATCACAAACCAATCAACTTGAGAGAAGCGCA 1260
Db 1201 AAAGAAAGGCTCCGAAGAAAGAGGAGATATCACAAACCAATCAACTTGAGAGAAGCGCA 1260
Qy 1261 GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAGACCC 1320
Db 1261 GCCCGATCTTTCTAAACAACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAGACCC 1320
Qy 1321 CCAGCTTCAGGACCTCGACATGATCTCACCTGTGTAGAGATCAAAAGAAAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTCGACATGATCTCACCTGTGTAGAGATCAAAAGAAAGAGCTTTGAT 1380
Qy 1381 GCTCCCAACACTTTCAAACCTCAAAAGGCCATGTTATCTCGTCAACAAAGGAATCGGAAA 1440
Db 1381 GCTCCCAACACTTTCAAACCTCAAAAGGCCATGTTATCTCGTCAACAAAGGAATCGGAAA 1440
Qy 1441 CTTTGAATCTGCTGGCTGTAAAGAAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
Db 1441 CTTTGAATCTGCTGGCTGTAAAGAAAAGAGCAACAAAGAGGGACGGCGGGAAGAGGA 1500
Qy 1501 GGACGAAGACGAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACACGAGGTT 1560
Db 1501 GGACGAAGACGAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACACGAGGTT 1560
Qy 1561 GAAGGAAGCGAGTGTCTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCTC 1620
Db 1561 GAAGGAAGCGAGTGTCTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCTC 1620
Qy 1621 CGAATCTCATCTGCTTTGGCTTCGGTATCAACGCTGAAAAACAAACAGAGATCTTCTTGC 1680
Db 1621 CGAATCTCATCTGCTTTGGCTTCGGTATCAACGCTGAAAAACAAACAGAGATCTTCTTGC 1680
Qy 1681 AGGTGATAGGACAAATGTGTAGACAGATAGAGAGCAAGCAAGAGATTTAGCATTTCC 1740
Db 1681 AGGTGATAGGACAAATGTGTAGACAGATAGAGAGCAAGCAAGAGATTTAGCATTTCC 1740
Qy 1741 TGGTCCGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGGAATCTCACTTTGTAG 1800
Db 1741 TGGTCCGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGGAATCTCACTTTGTAG 1800
Qy 1801 TGTCTGCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGTCTGCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAAAAACCAAGSAGGAGGTCCACTCTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGGAGGAAAAACCAAGSAGGAGGAGGTCCACTCTCTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTGAGAATGAGGCAACTTCTTATGTATCGATAATAAGATCAAGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAATGAGGCAACTTCTTATGTATCGATAATAAGATCAAGCTTTTGTACT 1980
Qy 1981 CTACTATCAAAAACTTATCAATAAAAAAGCTTTTGGCTTTTCTCC 2032
Db 1981 CTACTATCAAAAACTTATCAATAAAAAAGCTTTTGGCTTTTCTCC 2032

RESULT 2
US-10-228-806-1
; Sequence 1, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2032

; TYPE: DNA									
; ORGANISM: Arachis hypogaea									
US-10-228-806-1									
Query Match 100.0%; Score 2032; DB 14; Length 2032;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAG	60						
Db	1	AATAATCATATATATTCATCAATCATCTATATATAAGTAGTAGCAGGAGCAATGAGAGGAG	60						
Qy	61	GGTTTCTCCAGTCGATGCTGTGTGCTAGGGATCTTGTCTGCTGGCTTCTGCAACGCA	120						
Db	61	GGTTTCTCCAGTCGATGCTGTGTGCTAGGGATCTTGTCTGCTGGCTTCTGCAACGCA	120						
Qy	121	TGCCAAGTCATCACTTACCAGAAGAAACACAGAACCCCTCGCCACAGAGGTCCTCCA	180						
Db	121	TGCCAAGTCATCACTTACCAGAAGAAACACAGAACCCCTCGCCACAGAGGTCCTCCA	180						
Qy	181	GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240						
Db	181	GAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCTCGCTGCACAA	240						
Qy	241	GCTCGAGTAGATCTCTGTTGTGTATGATCTCTCGAGGACACACTGCGACACCAACCA	300						
Db	241	GCTCGAGTAGATCTCTGTTGTGTATGATCTCTCGAGGACACACTGCGACACCAACCA	300						
Qy	301	ACGTTCCCTCCAGGGGAGCGGACAGCTGGCCGCAACCCGCGAGACTACGATGATGACCG	360						
Db	301	ACGTTCCCTCCAGGGGAGCGGACAGCTGGCCGCAACCCGCGAGACTACGATGATGACCG	360						
Qy	361	CCGTCAACCCGGAAGAGAGAGGAGCGGCGATGGGACACAGCTGGAGAGCGTGA	420						
Db	361	CCGTCAACCCGGAAGAGAGAGGAGCGGCGATGGGACACAGCTGGAGAGCGTGA	420						
Qy	421	AAGAGAAGAGACTGAGAGCAACCAAGAGAAGATTGGAGGGGACCAAGTCAATGACGACC	480						
Db	421	AAGAGAAGAGACTGAGAGCAACCAAGAGAAGATTGGAGGGGACCAAGTCAATGACGACC	480						
Qy	481	ACGGAAATTAAGCCCGAAGGAGAGAGAGGAGAAACAAGTGGGGAAACACAGTAGGCA	540						
Db	481	ACGGAAATTAAGCCCGAAGGAGAGAGAGGAGAAACAAGTGGGGAAACACAGTAGGCA	540						
Qy	541	TGTGAGGGAAGAAACATCTCGGAAACAACCTTTTCTACTTCCCGTCAAGCGGTTTAGCAC	600						
Db	541	TGTGAGGGAAGAAACATCTCGGAAACAACCTTTTCTACTTCCCGTCAAGCGGTTTAGCAC	600						
Qy	601	CCGCTACGGGAACCAAAAACGGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAG	660						
Db	601	CCGCTACGGGAACCAAAAACGGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAG	660						
Qy	661	GCAGTTTCAGATCTCCAGAAATCAACGATATGTCAGATCGAGGCCAACCTACACTCT	720						
Db	661	GCAGTTTCAGATCTCCAGAAATCAACGATATGTCAGATCGAGGCCAACCTACACTCT	720						
Qy	721	TGTTCTTCCCAAGCAGCTGTGATTAACATCTTGTATTCACAGCAAGGGCAAGCCAC	780						
Db	721	TGTTCTTCCCAAGCAGCTGTGATTAACATCTTGTATTCACAGCAAGGGCAAGCCAC	780						
Qy	781	CGTGACCGTAGCAATATGGAAGAGGCTTTTAACTTTGACGAGGCCCATGCACT	840						
Db	781	CGTGACCGTAGCAATATGGAAGAGGCTTTTAACTTTGACGAGGCCCATGCACT	840						
Qy	841	CAGAAATCCCATCCGGTTTCATTTCTCTACATCTTGAACCGCCATGACACAGAACCTCAG	900						
Db	841	CAGAAATCCCATCCGGTTTCATTTCTCTACATCTTGAACCGCCATGACACAGAACCTCAG	900						
Qy	901	AGTAGCTTAAATCTCCATGCGCGTTAAACACACCCCGCAGTTTGAGGATTTCTTCCCGGC	960						
Db	901	AGTAGCTTAAATCTCCATGCGCGTTAAACACACCCCGCAGTTTGAGGATTTCTTCCCGGC	960						
Qy	961	GAGCAGCCGAGACCAATCACTCTTTCAGGCGCTTCAGCAGGAATACGTTTGGAGGCCGC	1020						
Db	961	GAGCAGCCGAGACCAATCACTCTTTCAGGCGCTTCAGCAGGAATACGTTTGGAGGCCGC	1020						

US-10-100-303A-5

; Sequence 5, Application US/10100303A
 ; Publication No. US20030202980A1
 ; GENERAL INFORMATION:

```
,
, APPLICANT: Caplan, et al.
, TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
,
, TITLE OF INVENTION: to Allergy
, FILE REFERENCE: 2002834-0166
, CURRENT APPLICATION NUMBER: US/10/100_303A
, CURRENT FILING DATE: 2002-03-18
, NUMBER OF SEQ ID NOS: 138
, SOFTWARE: PatentIn Ver. 2.1
,
```

; LENGTH: 2032

TYPE: DNA

ORGANISM: *Arachis hypogaea*, Prot/Nucleo-Arach 1

US-10-100-303A-5

Query Match	100.0%;	Score 2032;	DB 17;	Length 2032;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2032;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy 1 AATAATCATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGAG 60

D6 1 AATAATCATATATTCATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGAG 60

Qy 61 GGTTTCTCCACTGATGCTGTTGCTAGGATCCTTGTCTGGCTTCTAGTTTCTGCAACGCA 120

Db 61 GGTTTCTCCACTGATGCTGTTGCTAGGATCCTTGTCTGGCTTCTAGTTTCTGCAACGCA 120

Qy	121	TGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTGGCCCCAGAGGTCCTCCA	180
Dβ	121	TGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTGGCCCCAGAGGTCCTCCA	180

Qy	181	GAGTTGTCAA	CAGGAACCGGATGACTTTGAAGCAAAAGGCATGGAGTCTCGTGCACCAA	240
Dp	181	GAGTTGTCAA	CAGGAACCGGATGACTTTGAAGCAAAAGGCATGGAGTCTCGTGCACCAA	240

Qy	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCAACCA	300
Dβ	241	GCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGCACCAACCA	300

QY	301	ACGTTCCCTCCAGGGAGCGGACACGTGGCGGCCAACCCGGAGACTTACGATGATGACCG	360
db	301	ACGTTCCCTCCAGGGAGCGGACACTGGCGGCCAACCCGGAGACTTACGATGATGACCG	360

QY 361 CCGTCAACCCCGAAGAGAGGAGGAGCCGATGGGGACGAGTGGACCGAGGAGCGTGA 420

Dh 361 CCGTCAACCCCGAAGAGAGGAGGAGCCGATGGGGACGAGTGGACCGAGGAGCGTGA 420

Qy 421 AAGAGAAGAGACTTGGAGACAACTAAGAGAGAGATTGGAGGCCACCAAGTCATCAGCGAGCC 480

Qy 481 ACGGAAATTAAGGCCGAGGAAGAAGAGAAACAAGAGTGGGGAACAACAGGTAGGCA 540

Qy 541 TGTGAGGGAAGAAACATCTCGGAACAAACCCCTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600

Qy 601 CCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGACCAAAAGGTC AAG 660

Qy 661 GCAGTTTCAGAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAAACCTTAACACTCT 72

Qy 721 TGTTCCTTCCCAAGCAGCTGATGCTGATAACATCCTTGTATTATCCAGCAAGGGCAAGCCAC 78

Qy 1561 GAAGGAGGCGATGTGTTTCATATGCCAGAGCTCATCCAGTAGCCATCAAGGCTTCCTC 1620
Db 1561 GAAGGAGGCGATGTGTTTCATATGCCAGAGCTCATCCAGTAGCCATCAAGGCTTCCTC 1620
Qy 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACACAGAAATCTTCCCTGC 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACACAGAAATCTTCCCTGC 1680
Qy 1681 AGGTGATAAGACAAATGTGATAGACAGATAGAACCAAGCGAAGATTTAGCATTTCC 1740
Db 1681 AGGTGATAAGACAAATGTGATAGACAGATAGAACCAAGCGAAGATTTAGCATTTCC 1740
Qy 1741 TGGGTCCGGGTGAACAAAGTTCAGAGCTCATCAAAAAACAGAGAAATCTCACTTTGTGAG 1800
Db 1741 TGGGTCCGGGTGAACAAAGTTCAGAGCTCATCAAAAAACAGAGAAATCTCACTTTGTGAG 1800
Qy 1801 TGCTCGTCCCTCAATCTCAATCTCGTCTCCTCTGAGAAAGAGTCTCCTGAGAA 1860
Db 1801 TGCTCGTCCCTCAATCTCAATCTCGTCTCCTCTGAGAAAGAGTCTCCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGAGAGAAACCAAGGAGGAGGCTCCACTCTTCAATTTTGAAGGC 1920
Db 1861 AGAGGATCAAGAGAGAGAAACCAAGGAGGAGGCTCCACTCTTCAATTTTGAAGGC 1920
Qy 1921 TTTTAACTGAGAAATGGAGGCAACTGTGTATGATCGATAATAAGATCACGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAAATGGAGGCAACTGTGTATGATCGATAATAAGATCACGCTTTTGTACT 1980
Qy 1981 CTACTATCCAAAACTTATCAATTAATAAAAAAGTTTGTGGTTGTTTCTCC 2032
Db 1981 CTACTATCCAAAACTTATCAATTAATAAAAAAGTTTGTGGTTGTTTCTCC 2032

RESULT 5

US-10-728-051-1
; Sequence 1, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10/728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-728-051-1

Query Match 100.0%; Score 2032; DB 20; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGCAATGAGGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATAAGTAGTAGCAGGCAATGAGGGAG 60
Qy 61 GGTTCCTCCAGTGTGTTGCTAGGAGCTTGTCTCGCTTCACTTCTGCTCAACGCA 120
Db 61 GGTTCCTCCAGTGTGTTGCTAGGAGCTTGTCTCGCTTCACTTCTGCTCAACGCA 120
Qy 121 TGCCAAGTTCATCACCTTACCAGAGAAAAACAGAGAACCCCTGCGCCAGAGGTCCTCCA 180

Db 121 TGCCAAGTTCATCACCTTACCAGAGAAAAACAGAGAACCCCTGCGCCAGAGGTCCTCCA 180
Qy 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGCTGCACCAA 240
Db 181 GAGTTGTCAACAGGAACCGGATGACTTTGAAGCAAAAGGATCGAGTCTCGCTGCACCAA 240
Qy 241 GCTCGAGTATGATCTCTCGTTGTCTATATGATCTCTCGAGGACACACTTGGCACCAACCA 300
Db 241 GCTCGAGTATGATCTCTCGTTGTCTATATGATCTCTCGAGGACACACTTGGCACCAACCA 300
Qy 301 ACGTTCCTCCCTCCAGGGAGCGGACACGTGGCCGCGCAACCCGGAGACTAAGATGATGACCG 360
Db 301 ACGTTCCTCCCTCCAGGGAGCGGACACGTGGCCGCGCAACCCGGAGACTAAGATGATGACCG 360
Qy 361 CGGTCAACCCCGAAGAGAGAGAGGAGCGGATGGGACCGAGCTGGACCCGAGGGAGCGTGA 420
Db 361 CGGTCAACCCCGAAGAGAGAGAGGAGCGGATGGGACCGAGCTGGACCCGAGGGAGCGTGA 420
Qy 421 AAGAGAAAGAGACTGGAGACCAACCAAGAAAGATTGGAGGCGACCAAGTCTCATCAGCAGCC 480
Db 421 AAGAGAAAGAGACTGGAGACCAACCAAGAAAGATTGGAGGCGACCAAGTCTCATCAGCAGCC 480
Qy 481 ACGGAAAAATAAGGCCCGGAGGAGAGAAAGAGAGAAAGAGAGTGGGAAACACACAGGTAGCCA 540
Db 481 ACGGAAAAATAAGGCCCGGAGGAGAGAAAGAGAGAAAGAGAGTGGGAAACACACAGGTAGCCA 540
Qy 541 TGTGAGGAGAGAAACATCTCGGAAACACCTTTCTACTTCCGTCGAAGGGGTTTATGAC 600
Db 541 TGTGAGGAGAGAAACATCTCGGAAACACCTTTCTACTTCCGTCGAAGGGGTTTATGAC 600
Qy 601 CCGCTACGGGAAACAAAACGTTAGGATCGGGGCTCTGACAGAGGTTTGACCAAAAGTCAAG 660
Db 601 CCGCTACGGGAAACAAAACGTTAGGATCGGGGCTCTGACAGAGGTTTGACCAAAAGTCAAG 660
Qy 661 GCAGTTTCAGAAATCTCCAGAAATCAACCGTATTTGTCAGATCGAGGCCAAACCTTAACTCT 720
Db 661 GCAGTTTCAGAAATCTCCAGAAATCAACCGTATTTGTCAGATCGAGGCCAAACCTTAACTCT 720
Qy 721 TGTTCCTCCAGCAGCGCTGATCTGATAACCTTTGTTATCCACAGGCGCAAGCCAC 780
Db 721 TGTTCCTCCAGCAGCGCTGATCTGATAACCTTTGTTATCCAGCAAGGGCAAGCCAC 780
Qy 781 CGTGACGCTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCATGCAC 840
Db 781 CGTGACGCTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTGACGAGGCCATGCAC 840
Qy 841 CAGAAATCCCATCCGGTTTCAATTTCTTACATCTTTGAAACCGCCATGACAAACCTCAG 900
Db 841 CAGAAATCCCATCCGGTTTCAATTTCTTACATCTTTGAAACCGCCATGACAAACCTCAG 900
Qy 901 AGTAGCTTAAATCTCCATCCCGTTAACACACCCGGCCAGTTTGGAGATTTCTTCCCGGC 960
Db 901 AGTAGCTTAAATCTCCATGCCCTTAAACACACCCGGCCAGTTTGGAGATTTCTTCCCGGC 960
Qy 961 GAGCAGCCGAGACCAATCATCTTACTTTGACGGCTTTCAGCAGGAAATACGTTGGAGGCCG 1020
Db 961 GAGCAGCCGAGACCAATCATCTTACTTTGACGGCTTTCAGCAGGAAATACGTTGGAGGCCG 1020
Qy 1021 CTTCAATGCGGAATTCATAGATAACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA 1080
Db 1021 CTTCAATGCGGAATTCATAGATAACGAGGGTGTCTTTAGAGAGAAATGACAGAGGTGA 1080
Qy 1081 GCAAGAGGAGAGGGCGAGGCGATCGGAGTACTCGGAGTAGTGAGAACTAAGAGGT 1140
Db 1081 GCAAGAGGAGAGGGCGAGGCGATCGGAGTACTCGGAGTAGTGAGAACTAAGAGGT 1140
Qy 1141 GATAGTCAAAAGTCTCAAAAGAGGACGTTGAAGAACTTACTAAGCAGCTAAATCCGCTCTC 1200
Db 1141 GATAGTCAAAAGTCTCAAAAGAGGACGTTGAAGAACTTACTAAGCAGCTAAATCCGCTCTC 1200
Qy 1201 AAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCCCAATCAATTTGAGAGAGCGGA 1260
Db 1201 AAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCCCAATCAATTTGAGAGAGCGGA 1260

Db 1201 AAGAAAGGCTCCGAAGAAGAGGAGATATACCAACCCAACTCACTTTGAGGAAGGCGA 1260
Qy 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCC 1320
Db 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAAGAAACCC 1320
Qy 1321 CCAGCTTCAGGACCTGGACATGATGCTCCTCAGCTGTGTAGAGATCAAAAGAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGACATGATGCTCCTCAGCTGTGTAGAGATCAAAAGAGAGCTTTGAT 1380
Qy 1381 GCTCCCACTTCAACTCAAGGCCATCGTTTATCGTCTCGTCTCAACAAAGGAACTGGAAA 1440
Db 1381 GCTCCCACTTCAACTCAAGGCCATCGTTTATCGTCTCGTCTCAACAAAGGAACTGGAAA 1440
Qy 1441 CTTGAACTCGTGGCTGTAAAGAAAGACAAACAGAGGGGACGGCGGGAAGAGAGGA 1500
Db 1441 CTTGAACTCGTGGCTGTAAAGAAAGACAAACAGAGGGGACGGCGGGAAGAGAGGA 1500
Qy 1501 GGACGAACCAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
Db 1501 GGACGAACCAAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTT 1560
Qy 1561 GAAGGAAGCGATGTGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGGCTTCTC 1620
Db 1561 GAAGGAAGCGATGTGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGGCTTCTC 1620
Qy 1621 CGAACTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAAACAACAGAGAACTCTTCTGC 1680
Db 1621 CGAACTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAAACAACAGAGAACTCTTCTGC 1680
Qy 1681 AGGTGATAGACCAATGTGATAGACCATAGAGAGGAGCAAGGAGGATTTAGCATTTCC 1740
Db 1681 AGGTGATAGACCAATGTGATAGACCATAGAGAGGAGCAAGGAGGATTTAGCATTTCC 1740
Qy 1741 TGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800
Db 1741 TGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800
Qy 1801 TGCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCCTGAGAAAGAGTCTCTGAGAA 1860
Db 1801 TGCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCCTGAGAAAGAGTCTCTGAGAA 1860
Qy 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Db 1861 AGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1920
Qy 1921 TTTTAACTGAGAACTGAGGCAACTTGTATGTATCGATAATAAGATCAAGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAACTGAGGCAACTTGTATGTATCGATAATAAGATCAAGCTTTTGTACT 1980
Qy 1981 CTACTATCCAAAACTTATCAATAATAAAAAAGCTTTGTGCTTTCTCC 2032
Db 1981 CTACTATCCAAAACTTATCAATAATAAAAAAGCTTTGTGCTTTCTCC 2032

RESULT 6

US-10-899-551-1
; Sequence 1, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2032

; TYPE: DNA
; ORGANISM: species Arachis hypogea
US-10-899-551-1

Query Match 100.0%; Score 2032; DB 21; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGAGCAATAGAGGGAG 60
Db 1 AATAATCATATATATTCATCAATCATCTATATATAGTAGTAGCAGAGCAATAGAGGGAG 60
Qy 61 GGTTCCTCCACTGATGCTGTGTAGGGATCCTGTCTGCTTCACTTTCTGCAACGCA 120
Db 61 GGTTCCTCCACTGATGCTGTGTAGGGATCCTGTCTGCTTCACTTTCTGCAACGCA 120
Qy 121 TGCCAACTCATCACTTTACAGAGAAACAGAGAACCCCTCGCCAGAGGTCCTTCCA 180
Db 121 TGCCAACTCATCACTTTACAGAGAAACAGAGAACCCCTCGCCAGAGGTCCTTCCA 180
Qy 181 GAGTTGTCAACAGAAACCGGATGATCTTTGAAAGCAAAAGGCATGCGAGTCTCGTGCACAA 240
Db 181 GAGTTGTCAACAGAAACCGGATGATCTTTGAAAGCAAAAGGCATGCGAGTCTCGTGCACAA 240
Qy 241 GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA 300
Db 241 GCTCGAGTATGATCCTCGTTGTGTCTATGATCTCGAGGACACACTGGCACCAACCA 300
Qy 301 ACCTTCCCTCCAGGGAGCGGACACGTCGCGCAACCCGAGACTACGATGATGACCG 360
Db 301 ACCTTCCCTCCAGGGAGCGGACACGTCGCGCAACCCGAGACTACGATGATGACCG 360
Qy 361 CCCTCAACCCCGAAGAGAGGAGCGGATGAGGACCAAGTGGACCGAGGAGCGTGA 420
Db 361 CCCTCAACCCCGAAGAGAGGAGCGGATGAGGACCAAGTGGACCGAGGAGCGTGA 420
Qy 421 AAGAGAGAGACTGGAGACAAACAGAGAGAAAGATTGGAGGCGCAAGTCAACAGAGCC 480
Db 421 AAGAGAGAGACTGGAGACAAACAGAGAGAAAGATTGGAGGCGCAAGTCAACAGAGCC 480
Qy 481 ACCGAAATTAAGCCCGAAG 540
Db 481 ACCGAAATTAAGCCCGAAG 540
Qy 541 TGTGAGGAGAGAAACATCTCGGAAACACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Db 541 TGTGAGGAGAGAAACATCTCGGAAACACCTTTCTACTTCCCGTCAAGGCGGTTTAGCAC 600
Qy 601 CCCTACCGGAAACCAAAACCGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCAG 660
Db 601 CCCTACCGGAAACCAAAACCGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCAG 660
Qy 661 GCAGTTTCAGAACTCCAGAAATCACGTTATTTGTCAGATCGAGGCCAAACCTAACACTCT 720
Db 661 GCAGTTTCAGAACTCCAGAAATCACGTTATTTGTCAGATCGAGGCCAAACCTAACACTCT 720
Qy 721 TGTTCCTCCCAAGCAGCTGATCTGATAACATCTTTGTTATCCAGAGGCGCAAGCCAC 780
Db 721 TGTTCCTCCCAAGCAGCTGATCTGATAACATCTTTGTTATCCAGAGGCGCAAGCCAC 780
Qy 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACT 840
Db 781 CGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACT 840
Qy 841 CAGAACTCCATCCGTTTCAATTTCTTACATCTTTGAAACCGGCATGACCAACAGACCTCAG 900
Db 841 CAGAACTCCATCCGTTTCAATTTCTTACATCTTTGAAACCGGCATGACCAACAGACCTCAG 900
Qy 901 AGTAGCTAAAATCTCCATGCCGCTTAAACACACCCCGGCACTTTGAGGATTTCTTCCGGC 960
Db 901 AGTAGCTAAAATCTCCATGCCGCTTAAACACACCCCGGCACTTTGAGGATTTCTTCCGGC 960
Qy 961 GAGCAGCCGAGACCAATCATCTTACTTGTGAGGCGTTCAGCAGGAATACGTTGAGGCGCGC 1020

[illegible]

```

US-10-958-324-5
; Sequence 5, Application US/10958324
; Publication NO. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFI N'DA
; APPLICANT: VIQUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; TITLE OF INVENTION: TRANSGENIC PEANUT SEEDS
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,955
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-958-324-5

```


QY 661 GCAGTTTTCAGAAATCTCCAGAAATCCCGTATTGTGAGATCGAGGCGCAACCTTAACACTCT 720
Db 661 GCAGTTTTCAGAAATCTCCAGAAATCCCGTATTGTGAGATCGAGGCGCAACCTTAACACTCT 720
QY 721 TGTCTTCTCCAAAGCAGCTGATGCTGATAAACAATCTTGTATCCAGCAAGGCGCAAGCCAC 780
Db 721 TGTCTTCTCCAAAGCAGCTGATGCTGATAAACAATCTTGTATCCAGCAAGGCGCAAGCCAC 780
QY 781 CGTAGCCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGAAGGCGCATGCACT 840
Db 781 CGTAGCCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGAAGGCGCATGCACT 840
QY 841 CAGAATCCCATCCGCTTTCATTTCTCAATCTTGAACCGCATGACCAACCAAGAACTCTCAG 900
Db 841 CAGAATCCCATCCGCTTTCATTTCTCAATCTTGAACCGCATGACCAACCAAGAACTCTCAG 900
QY 901 AGTAGCTAAATCTCCATGCCGTTTAAACACCCCGGCCAGTTTGAAGGATTTCTTCCCGGC 960
Db 901 AGTAGCTAAATCTCCATGCCGTTTAAACACCCCGGCCAGTTTGAAGGATTTCTTCCCGGC 960
QY 961 GAGGAGCCGAGACCAATCATCTACTTGCAGGGCTTCAGCGGAATAGCTTGGAGGCCGC 1020
Db 961 GAGGAGCCGAGACCAATCATCTACTTGCAGGGCTTCAGCGGAATAGCTTGGAGGCCGC 1020
QY 1021 CTTCAATCGGAATTCATAGATACGGAGGCTGCTGTTAGAGAGAAATGCGAGAGTGA 1080
Db 1021 CTTCAATCGGAATTCATAGATACGGAGGCTGCTGTTAGAGAGAAATGCGAGAGTGA 1080
QY 1081 GCAAGAGAGAGAGGCGAGCGGATGAGTACTCGGAGTAGTGAGAAACAATGAAGAGT 1140
Db 1081 GCAAGAGAGAGAGGCGAGCGGATGAGTACTCGGAGTAGTGAGAAACAATGAAGAGT 1140
QY 1141 GATAGTCAAAGTGTCAAAGGAGCAGTGTGAAGAACTTATCAAGCAAGCTAAATCGTCTC 1200
Db 1141 GATAGTCAAAGTGTCAAAGGAGCAGTGTGAAGAACTTATCAAGCAAGCTAAATCGTCTC 1200
QY 1201 AAAGAAAGGCTCCGAGAGAGGAGGAGATATACCAACCCCAATCACTTGAAGAGGCGA 1260
Db 1201 AAAGAAAGGCTCCGAGAGAGGAGGAGATATACCAACCCCAATCACTTGAAGAGGCGA 1260
QY 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTATTTGAGGTGAAGCCAGACAAGAAAGCC 1320
Db 1261 GCCCGATCTTTCTAAACAATTTGGGAAGTATTTGAGGTGAAGCCAGACAAGAAAGCC 1320
QY 1321 CCAGCTTCAGGACCTGGAATGATGCTCACTGTGTAGAGATCAAGAAAGAGCTTTGAT 1380
Db 1321 CCAGCTTCAGGACCTGGAATGATGCTCACTGTGTAGAGATCAAGAAAGAGCTTTGAT 1380
QY 1381 GCTCCCACTTCACTCAAAGGCCATGGTTATCGTGTGTGATCAAGAAAGCTGGAAA 1440
Db 1381 GCTCCCACTTCACTCAAAGGCCATGGTTATCGTGTGTGATCAAGAAAGCTGGAAA 1440
QY 1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACAAGAGGCGGCGGGAAGAGAGGA 1500
Db 1441 CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACAAGAGGCGGCGGGAAGAGAGGA 1500
QY 1501 GGAAGAGAGAGAGAGAGAGAGAGTAACAGAGAGTGGCTAGGTAACAAGAGAGTT 1560
Db 1501 GGAAGAGAGAGAGAGAGAGAGAGTAACAGAGAGTGGCTAGGTAACAAGAGAGTT 1560
QY 1561 GAAGGAGGCGATGTGTTTCATCATGCGCAGAGCTCATCCAGTAGCCATCAAGCTTCTC 1620
Db 1561 GAAGGAGGCGATGTGTTTCATCATGCGCAGAGCTCATCCAGTAGCCATCAAGCTTCTC 1620
QY 1621 CGAACTCCATCTGCTTGGCTTCGCTATCAACGCTGAAAAACAACAAGAACTTCTTTCG 1680
Db 1621 CGAACTCCATCTGCTTGGCTTCGCTATCAACGCTGAAAAACAACAAGAACTTCTTTCG 1680
QY 1681 AGGTGATAGGACCAATGTGATAGACCAATAGAGAGCAAGCGAAGGATTTAGCATTTCC 1740
Db 1681 AGGTGATAGGACCAATGTGATAGACCAATAGAGAGCAAGCGAAGGATTTAGCATTTCC 1740
QY 1741 TGGTCCGGTGAACAAGTTGAGAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800

Db 1741 TGGTCCGGTGAACAAGTTGAGAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAG 1800
QY 1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAAGAGTCTCTGTGAAA 1860
Db 1801 TGCTCGTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAAGAGTCTCTGTGAAA 1860
QY 1861 AGAGGATCAAG 1920
Db 1861 AGAGGATCAAG 1920
QY 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGATCGATAAATAAGATACGCTTTTGTACT 1980
Db 1921 TTTTAACTGAGAAATGGAGGCAACTTGTATGATCGATAAATAAGATACGCTTTTGTACT 1980
QY 1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGAGTGTGCGTTGTCTTC 2032
Db 1981 CTACTATCCAAAACTTATCAATAAATAAATAAAGAGTGTGCGTTGTCTTC 2032
RESULT 8
US-10-100-303A-6
; Sequence 6, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100-303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-6
Query Match 86.2%; Score 1752.4; DB 17; Length 1949;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1894; Conservative 0; Mismatches 46; Indels 54; Gaps 6;
QY 48 CAATGAGAGGAGGGTTTCTCACTGATGCTGTGTGAGATCCTTGTCTGCTGCTCAG 107
Db 1 CAATGAGAGGAGGGTTTCTCACTGATGCTGTGTGAGATCCTTGTCTGCTGCTCAG 60
QY 108 TTTCTGCAACGCGATGCGCAAGTCACTTACCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 167
Db 61 TTTCTGCAACGCGATGCGCAAGTCACTTACCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 114
QY 168 AGAGGTGCTCCAGAGTTGCAACAGGAAACCGGATGACTTGAAGCAAAAGCATGCGAGT 227
Db 115 AGAGGTGCTCCAGAGTTGCAACAGGAAACCGGATGACTTGAAGCAAAAGCATGCGAGT 174
QY 228 CTGCTGCAACAGCTCGAGTATGATCCTGTTGTGTGTATGATCCTCGAGGACACACTG 287
Db 175 CTGCTGCAACAGCTCGAGTATGATCCTGTTGTGTGTATG-----ACACTG 222
QY 288 GCACCAACCAACAGCTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGAGACT 347
Db 223 GCGCCACCAACCAACAGCTTCCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGAGACT 282
QY 348 ACAGTATGATGCGCGCTCAACCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407
Db 283 ACAGTATGATGCGCGCTCAACCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 342
QY 408 CGAGGAGAGCTGAAAG 467
Db 343 CGAGGAGAGCTGAAAG 402
QY 468 GTCATCAGCAGCCCAACGAGAAATTAAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 527

TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
CLASSES OF SOYBEAN SEED
PROTEIN GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/684,651
FILING DATE: 14-Oct-2003
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/758,652
FILING DATE: 11-Jan-2001
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14, 1996

ATTORNEY/AGENT INFORMATION:
NAME: LYNN M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-684-651-3
Query Match 20.0%; Score 406.8; DB 18; Length 1320;
Best Local Similarity 62.1%; Pred. No. 2.4e-115;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;
562 GAAACACCTTTCTACTT---CCGCTCAAGCGGTTTAGCACCGCTACGGGAACCAAAA 618
90 GAATAACCCCTTCTACTTTAGAACTCTAACAGCTCTAACAGCTTCCAAACTCTCTTTGAGAACCAAAA 149
619 CGGTAGGATCGGCTCCGTCAGAGGTTTGACCAAGGTTCAAGGAGTTTCAGAAATCTCCA 678
150 CGTTGCGATTCGTTCTCTCCAGAGATTCAACAAACGCTCCCAACCTTGAGAACCTTCG 209
679 GAATCACCGTATTGTCAGATCGAGGCAACCTTAACACTCTGTTCTTCCCAAGCAGC 738
210 AGACTACCGGATTTGTCAGTTTCAGTCAAAACCCCAACAACTCTTCTCCCCACCATGC 269
739 TGATGCTGATAACATCTTTGTTATCCAGCAAGGCAAGCCACCGTGACCGTAGCAATGG 798
270 TGACGCGGATTTCTCTCTTTGTTCTTAGCGGGAGAGCCATCTTACCTTTGGTGAACAA 329
799 CAATAACAGAAGAGCTTTATCTTGACGGGCGCATCTCAGATCCCATCCGTTT 858
330 CGACGACAGAGACTCTTCAACACCTTCAACCTGCGGATGCCAGAGATCCAGCTGGAAC 389
859 CATTTCTTACATCTTGAACCCCGCATGCAACACAGAACTCTCAGATGAGTAAATCTCCAT 918
390 CACTTACTATTGTTAACTTCAAGCCATCAGACCAACAGATCTCAAAATATCAAACTTGCAT 449
919 GCGCGTTAAACACACCCGCGCAATTTGAGGATTTCTTCCGGGAGAGCGGAGCCAAATC 978
450 ACCCGTCAACAAACCTGGCAGATATGATGATTTCTTCTTATCTAGCACTCAAGCCCAACA 509

979 ATCTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGGCGCCTTCAATCGGAATTCAA 1038
510 GTCTACTTGCAGGGCTTCAGCCATATATTTAGAGACCTCTTCCATAGCGAATTCGA 569
1039 TGAGATACGAGGGTCTGTTTGAAGAGAAATGAGAGGTTGACCAAGAGAGAGAGGCA 1098
570 GGAGATAAACAGGGTTTGTGTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 621
1099 GAGGCGATGGAGTACTCGGAGTAGTGAGAAATGAGAGGAGTATAGTCAAGTGTCAAA 1158
622 -----GGAGTAGTCTGGAATCTCAAA 644
1159 GGAGCACGTTTGAAGAACTTACTAAGCAACGCTTAATCCGCTCAAGAAAGGCTCCGAAGA 1218
645 GGACAAATTCGGCAACTGAGCAGAGCTGCAATCTAGTTCAAGGAAACCAATTCCTC 704
1219 AGAGGAGATATCACCAACCAATCACTTTGAGAGAGGAGGAGGAGGAGGAGGAGGAG 1278
705 CGAAGATG-----AACCATTTCACTTGAAGCGGCAACCCCATCTTATTTCCAAACA 755
1279 CTTTGGGAAGTTATTTGAGTGAAGCAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAGAGAGAGAGAGAGAGAGAGAG 812
1339 CATGATGCTCACCTGTGTAGAGATCAAGAAAGAGAGCTTTGATGCTCCCACTTCAACTC 1398
813 TATCTTCTCAGTTCTGTGATATCAAGAGAGAGCTCTTCTTCAACCACTTCAATTC 872
1399 AAAGGCAATGTTTATCGTCTCAACAAAGAGAGCTTTGAGAACTTGAACCTTGGCTGT 1458
873 AAAGGCAATGATGATGATGATTAATGAAGAGAGATGCAAAACATTTGAACCTTGTG 932
1459 AAGAAAGAGCAACACAGAGGAGGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1518
933 TAAAGAAACAAACAG 975
1519 GAGGGAAGGTAACAGAGAGGTTGCTAGGTACACAGCGAGGTTTGAAGAGAGCGATGTGT 1578
976 -----GAGTGCAAGGTACAGAGCTGAATTTGCTGCAAGACGATGTATT 1019
1579 CATCATGCGAGAGCTCATCCAGTAGCATCAACGCTTCTCCGAACTCCATCTGCTTGG 1638
1020 TGTAAATTCAGCAGAGCTTATCCATTTGCTGCTCAACGCTACCTCAACCTCAATTTCTT 1079
1639 CTTTGGTATCAACGCTGAAACCAACACAGAACTTCTTTCAGAGTGTATAGGACAAATGT 1698
1080 TTTTGGTATCAATGCTGAGAAACCAACAGAGAACTTCTTTCAGAGGCGGAGAGAGAGAG 1139
1699 GATAGACCAATAGAGAGAGCAAGGAGATTTAGCATTTCCCTGGGTCCGGTGAACAAGT 1758
1140 GGTAGGAGATAGAGAGAGCAAGTGCAGAGCTTGGTTCCTGGGTCTGCACAGATGT 1199
1759 TGAGAGCTCATCAAAACCAAGAGAGATCTCACTTTTGTGAGTGTGCTGCTCTCA 1812
1200 TGAGAGGCTATTAAAGAGAGAGAGGAGATCCTACTTTGTTGATGCTCAGGCTCA 1253

RESULT 14
US-10-757-074-3
Sequence 3, Application US/10757074
Publication No. US20040139502A1
GENERAL INFORMATION:
APPLICANT: ANTHONY JOHN KINNEY
TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
CLASSES OF SOYBEAN SEED
PROTEIN GENES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON


```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/757,155
; FILING DATE: 14-Jan-2004
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
us-10-757-155-3

Query Match      20.0%; Score 406.8; DB 19; Length 1320;
Best Local Similarity 62.1%; Pred. No. 2.4e-115;
Matches 779; Conservative 0; Mismatches 382; Indels 93; Gaps 5;

Qy 562 GAACAACCCCTTTCTACTT---CCCGTCAGGCGGTTTAGCACCGCTACGGGACCAAAAA 618
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 619 CGGTAGGATCGGGTCTCTGACAGGTTTGACAAAGGTCAAGGAGTTTCAGATCTCCA 678
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 150 CGTTCCGATTCGTCTCTCCAGAGATTCAACAAACGCTCCCAACAACCTTGAGAACCTTCG 209
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 679 GAATCACCGTATCTGACAGTCGAGGCAACACCTTAACACTCTTGTCTTCCCAAGCAGC 738
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 210 AGATACCGGATTCGAGTTTCTAGTCAAAACCAACCAACATCTCTTCCCCCACCATGC 269
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 739 TGATGCTGATAAATCTTGTATTCCAGCAAGGCAACCGCTGACCTAGCCATAGCAATGG 798
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 270 TGACGCCGATTTCTCTCTTGTCTTAGCGGAGAGCCATCTTACCTTGTGTAACA 329
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 799 CAATAACAGAAAGCTTTTAATCTTGACGAGGCGCATGCACTCAGAAATCCCATCCGGTTT 858
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 330 CGACGACAGAGACTCTCTACAACTTCAACCTTCCCTGGCGATGCCAGAGAAATCCGAGTGAAC 389
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 859 CATTTCTTACATCTTGAACGCCATGACAAACCAACCTCAGAGTAGCTAAATCTCCAT 918
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 390 CACTTACTATTGTTTAACTTCAACCTTCAAGCAACCAATCTCAAAATATCAAACTTGCCAT 449
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 919 GCCCGTTAAACACACCCGGCCAGTTTGAGGATTTCTTCCCGCGAGCAGCCGAGACCAATC 978
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 450 ACCCGTCAACAAACCTTGGCAGATATGATGATTTCTTCTATCTAGCACTCAAGCCCAACA 509
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 979 ATCTACTTTCAGGGCTTTCAGCAGGAATACCTTGGAGGCGCCCTTCAATGCGGAATTCAA 1038
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 510 GTCTACTTTCAGAGGCTTTCAGCCATAATTTCTAGAGACCTCTCTTCCATAGCGGAATTCGA 569
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1039 TGAGATACGAGGGTGTCTTTAGAGAGAAATGACAGGAGGTGAGCAAGAGGAGAGAGGCA 1098
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 570 GGAGATAAACAGGGGTTTGTGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1099 GAGGCGATGGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTAGTAGTCAAGTGTCAAA 1158
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1159 GGAGCAGCTTGAAGACTTACTAAGCAGCGCTAAATCCGTCTCAAGAAAGAGGCTCCGAAGA 1218
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 645 GGAACAAATTCGGCAACTGAGCAGACGCTGCCAAATCTAGTTCAAGGAAACCAATTTCTCTC 704
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 1219 AGAGGGAGATATCACCAACCCCAATCAACTTGAGAGAAGCGAGCCCGATCTTTTCTAACAA 1278
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 705 CGAAGATG-----AACCATTTCACTTGAGAAGCGCAACCCCATCTTATTTCCAAACA 755
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1279 CTTTGGGAAGTTTATTTGAGGTGAAGCCAGACAAGAAAGACCCCGAGCTTCAGGACCTGGA 1338
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 756 CTTTGGAAAGTTCTTTGAGATCACCCCTG---AGAAAAAACCCACAGCTTCGGGACTTGA 812
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1339 CATGATGCTCACCTGTGTAGAGATCAAGAAGAGCTTTGATGCTCCCACTTCACACTC 1398
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 813 TATCTTCTTCAGTTCTGTGGATATCAACGAAGAGCTCTTCTTACCACACTTCAATTC 872
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1399 AAAGGCCATGTTTATCGTCTCAACAAAGGAACTGGAAACCTTGAACCTCGTGGCTGT 1458
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 873 AAAGGCCATAGTGTAGTACTAGTATTGAAGGAGATGCAAAACATTGAACTTCTTGGCAT 932
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1459 AGAAAAAGAGCAACAAAGAGGGGACGGCGGGGAAGAAAGAGGAGGAGCAAGACGAAGA 1518
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 933 TAAAGAAACAAACAGAGAGCAGAAACAGGAAGAGGAACCTTTG----- 975
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1519 GAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGAGAGGTTGAAGGAAGGCGATGT 1578
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 976 -----GAAAGTCAAAGGTACAGAGCTGAATTTCTTGAAGACCATGTAAT 1019
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1579 CATCATGCCAGCAGCTCATCCAGTAGCCCATCAACGCTTCTCCGAACTCCCATCTGCTTGG 1638
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1020 TGTAAATCCAGCAGCTTATCCATTTGTCGTCACGCTACTCTCAAACTCAATTTCTTGC 1079
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1639 CTTTCGGTATCAACGCTGAAACAAACACAGAAATCTTCTTGCAGGTGATAAGGACAATGT 1698
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1080 TTTTGGTATCAATGCTGAGAAACAAACAGAGGAATCTTCTTGCAGGCGAGAAAGACAATGT 1139
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1699 GATAGACCAAGTAGAGAAAGCAAGCGAGGATTTAGCATTTCCCTGGGTCCGGTCAACAAGT 1758
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1140 GGTAAAGGAGATAGAAGACAAAGTCAGGAGCTTTCGTTTCCCTTGGGTCTGCACAAGATGT 1199
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1759 TGAGAAAGCTCATCAAAAAACAGAAAGGAATCTCACTTTGTGAGTGTCTCTCTCA 1812
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1200 TGAGAGGCTATTAAAGAACGAGAGGGAATCTTACTTTGTGTGCTCAGGCTCA 1253
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: August 24, 2005, 09:47:11
Job time : 1238.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 115.082 Seconds
(without alignments)
13658.075 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3562

Sequence: 1 aataacatcatatattcatc.....cggttgctggtgtttctcc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124355_29200/app_query.fasta_1.4757
-DB=A Geneseq 16Dec04 -OFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN_1_1224 @runat_23082005_124355_29200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq_16Dec04:*
1: Geneseqp19808:*
2: Geneseqp19908:*
3: Geneseqp20008:*
4: Geneseqp20018:*
5: Geneseqp20028:*
6: Geneseqp20038:*
7: Geneseqp2003bs:*
8: Geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	92.3	626	2	AAY15244 Peanut al
2	3286	92.3	626	2	AAY15244 Peanut al
3	3286	92.3	626	6	AAY25657 Peanut al
4	3286	92.3	626	7	ABUS2412 Peanut al
5	3282	92.1	626	4	ADG34906 Peanut al
6	3282	92.1	626	6	AAY4706 Anaphylac
7	3282	92.1	626	6	ABUS2570 Peanut Ar
8	3280	92.1	626	6	ABUS2568 Peanut Ar
9	3280	92.1	626	6	ABUS2571 Peanut Ar
10	3280	92.1	626	6	ABUS2574 Peanut Ar

11	3279	92.1	626	8	ADM12053 Arachis h
12	3278	92.0	626	6	ABUS2569 Peanut Ar
13	3278	92.0	626	6	ABUS2572 Peanut al
14	3274	91.9	626	2	AAY22150 Peanut al
15	3269.5	91.8	625	7	ADG27464 Peanut Ar
16	3251	91.3	626	8	ADO38271 Peanut al
17	3052	85.7	634	3	AB33599 Modified
18	3052	85.7	634	4	AAY4709 Modified
19	3052	85.7	634	4	AAY4709 Modified
20	3041	85.4	614	2	AAY22149 Peanut al
21	3041	85.4	614	2	AAY22149 Peanut al
22	3041	85.4	614	6	ABUS2413 Arachis h
23	3041	85.4	614	7	ADG27465 Peanut Ar
24	2130	59.8	415	2	AAY40913 Ara h 1 a
25	1281	36.0	605	2	AAY62838 Glycine m
26	1281	36.0	605	7	ADG27564 Soybean B
27	1276	35.8	605	7	ADG27564 Soybean B
28	1190	33.4	559	5	ABG71270 Glycine m
29	1180.5	33.1	543	5	ABG71271 Glycine m
30	1165	32.7	417	5	ABG71269 Glycine m
31	1157	32.5	417	6	ABP97239 Mature be
32	1157	32.5	425	6	ABP97238 FLAG-tag
33	1156	32.5	439	6	ABP97235 Glycine m
34	1117.5	31.4	390	6	ABP97236 Glycine m
35	865.5	24.3	666	2	AAY62829 Macadamia
36	863.5	24.2	625	2	AAY62830 Macadamia
37	849.5	23.8	666	2	AAY62828 Macadamia
38	811	22.8	566	2	AAR20181 Sequence
39	791	22.2	590	2	AAY62832 Gossypium
40	764	21.4	371	2	AAY40914 Phaseolin
41	737.5	20.7	525	2	AAY62831 Theobroma
42	654	18.4	366	6	ABUS2503 Soybean B
43	629.5	17.7	593	2	AAY62835 Zea mays
44	626	17.6	573	7	ADL18493 Maize glo
45	614	17.2	637	2	AAY62837 Hordeum v

ALIGNMENTS

RESULT 1

AAY15244

ID AAY15244 standard; protein; 626 AA.

AC AAY15244;

XX

DT 17-OCT-2003 (revised)

DT 09-NOV-1999 (first entry)

XX

DE Peanut allergen, Ara h 1, amino acid sequence.

XX

KW allergy; immune response; transgenic; allergen; epitope;

KW immunoglobulin E; Ig E; binding site; peanut.

XX

OS Arachis hypogaea.

XX

PN WO938978-A1.

XX

PD 05-AUG-1999.

XX

PF 29-JAN-1999; 99WO-US002031.

XX

PR 31-JAN-1998; 98US-0073283P.

PR

PR 13-FEB-1998; 98US-0074590P.

PR

PR 13-FEB-1998; 98US-0074624P.

PR

PR 13-FEB-1998; 98US-0074633P.

XX

PA (UYAR-) UNIV ARKANSAS.

PA

PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

XX

XX (SOSI/) SOSIN H.

PI

PI Sosin H, Bannan GA, Burks AW, Sampson HA;

DR WPI; 1999-479189/40.
 DR N-PSDB; AA206382.
 XX Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 XX Disclosure; Page 35-37; 46pp; English.
 XX This is the amino acid sequence of the Ara h 1 protein from Arachis
 CC hypogaea. The Ara h 1 protein has 23 IgE (Immunoglobulin E) binding
 CC epitopes, four of which are immunodominant (AA15247, AA15249, AA15250
 CC and AA15263). By modifying the IgE binding sites the ability of the
 CC allergen to provoke an immune response is downregulated. The epitopes of
 CC the IgE binding sites can therefore be modified in genetically engineered
 CC plants and animals to elicit less of an allergic response. (Updated on 17
 CC -OCT-2003 to standardise OS field)
 XX Sequence 626 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.4e-304 Length: 626
 Score: 3286.00 Matches: 626
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.25% Indels: 0
 DB: 2 Gaps: 0
 US-10-728-323-1 (1-2032) x AA151244 (1-626)
 QY 50 ATGAGAGGAGGGTTTCTCACTGATGCTGTGTGCTAGGATCTGTGCTGCTTCAGTT 109
 DB 1 MetArgLysArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACCATGCCAAGTCATCACCTTACCAGAGAGAAAACAGAGAACCCCTGCCGCCAG 169
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCCAGATGTCTCAACGAGAACCGGATGACTTCAAGCAAAAGCATGCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGTCGAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTGCG 289
 DB 61 ArgCysThrLysLeuGluLysArgProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGAGACGTCGCGCCCAACCCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGACCGACTGCGACCG 409
 DB 101 AspAspAspArgGlnProArgGlnGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAGAGAGAGAGACTGGAGACAAACAGAGAGATGGAGGCGACCAAGT 469
 DB 121 ArgGluArgGluGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
 QY 470 CATCAGCAGCCAGGAAATAAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGG 589
 DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTAGCCCGCTACGGGAACCAACCGTAGATCCGGGTCTCGAGAGGTTTGAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGTCAAGCGAGTTTCCAGATCTCCAGATCACCGTATTGTGCAGATCCAGGCCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTCTTCTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGAA 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCACCGCTGACCGTAGCAAAATGGCAATAACAGAAAAGAGCTTTAATCTTCA 829
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCATGCACTCAGATCCCATCCGTTTCTATCTTCTATCTTCTTGAACCGCATCACAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTCAGAT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTTCCCGGAGCAGCGCAGACCAATCATCTTCTTCTGAGGGCTTCAGCAGGAATACG 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGCGCGCTTCAATGCCGMAATTCAATGAGATACCGAGGGTCTCTTGAACAGAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGGAGTGAGCAAGAGAGAGAGGCGAGGCGATCGAGTACTCGGAGTAGTCAGAAC 1129
 DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
 QY 1130 AATGAGGAGTGATGATCAAAAGTGTCAAAGGAGCACGTTGAAGAACTTACTAAGCACGCT 1189
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGCTCAAGAAAGGCTCCGAAAGAGGAGGAGATATCAACCAACCCCAATCACTTG 1249
 DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAAGCGAGCGCCGATCTTTTAAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTTGAACATGATGCTCCTCCTGTGTAGAGATCAAAG 1369
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440
 QY 1370 GGAGCTTTGATGCTCCACACTCAACTCAAGGCCATGCTTATCGTCTGCTCAACAAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
 QY 1430 GGAATCGAAACCTTGAACTCTGCTGCTGTAAGAAAAGACCAACACAGAGGGGACGGCG 1489
 DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
 QY 1490 GAAGAGAGGAGGAG 1549
 DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGTTGAAGAGAGCGATGCTGTTTCATGCTGCGAGCAGCTCATCCAGTAGCCCATC 1609
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCTTTCAGGTGATAGGACAAATGTGATAGACCCAGATAGAGACAGACAGAGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTGGGTGAACAAAGTTGAGAGAGCTCATCAAAACACAGAGGAATCT 1789
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTGTGCTGCTCAATCTCAATCTCAATCTCGCTCGTCTCTCTGAGAAAGAG 1849

Db 581 HiePheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
Qy 1850 TCTCTGAGAAAGGATCAAGAGGAGGAAACCAAGAGGAGGATCCACTCTTCA 1909
Db 601 SerProGluLysGluArgGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
Qy 1910 ATTTTGAAGGCTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 2
ID AAY25657
XX AAY25657 standard; protein; 626 AA.
AC AAY25657;
XX 17-OCT-2003 (revised)
DT 30-SEP-1999 (first entry)
DE Peanut allergen 1168391 Ara h I protein fragment.
XX Major histocompatibility complex; class II; desensitising; human;
KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; peanut.
XX Arachis hypogaea.
OS
XX
XX WO9934826-A1.
XX 15-JUL-1999.
XX 11-JAN-1999; 99WO-GB000080.
XX 09-JAN-1998; 98GB-00000445.
PR 21-SEP-1998; 98GB-00020474.
XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX Larche M, Kay AB;
PI WPI; 1999-458255/38.
XX Desensitizing patients to polypeptide allergens.
PT
PS Example 6; Page 70-71; 117pp; English.
XX

This invention describes a novel method of desensitizing a patient to a polypeptide allergen and comprises administering to the patient a peptide derived from the allergen where restriction to a MHC Class II molecule possessed by the patient can be demonstrated for the peptide and the peptide is able to induce a late phase response in an individual who possesses the MHC Class II molecule. The methods can be used for desensitising patients to allergens present in e.g. grass, tree and weed (including ragweed) pollens, fungi and moulds, foods, stinging insects, the chironomidae (non-biting midges), spiders and mites, housefly, fruit fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee, non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig, sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to produce immunological vaccines which may be used to prevent and/or treat conditions involving hypersensitivity to allergens. This sequence represents a peanut (Arachis hypogaea) allergen 1168391 Ara h I. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 626 AA;
SQ
Alignment Scores:
Pred. No.: 1-4e-304 Length: 626
Score: 3286.00 Matches: 626

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-1 (1-2032) x AAY25657 (1-626)
Qy 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCTTCTCTGGGTTGAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
Qy 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAAGAAACAGAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProGlyGlnLysThrGlnAsnProCysAlaGln 40
Qy 170 AGGTGCTCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
Qy 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTGAGACACATGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
Qy 290 ACCACCAACCAACGTTCCCTCTCAGGGAGCGACACGTCGGCCGCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
Qy 350 GATGATGACCGCGCTCAACCCCGAAGAGAGGAGGCGCGATGGGGACCGAGCTGGACCG 409
Db 101 AspAspArgArgGlnProArgGluGluGlyArgTyrGlyProAlaGlyPro 120
Qy 410 AGGGAGCTGAAAGAGAGAGAACTGGAGACAAACAGAGAGAGATGGAGGGGACCAAGT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
Qy 470 CATCAGACCCACGGAATAAGGCCGGAAGAGAGAGAGAGAAACAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
Qy 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCGCTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
Qy 590 CGGTTTAGACCCGCTACGGGAACCAAAACGCTAGATCCGGGTCCTGAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
Qy 650 CAAAGTCAAGGCAGTTTCCAGAAATCTCCAGAAATCACCCTATTGTGCAGATCGAGGCCAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
Qy 710 CCTAACACTCTTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
Qy 770 GGGCAAGCCACCGTAGCCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
Qy 830 GGCCATGCACTCAGAAATCCCATCCGTTTCTTCTCTACATCTTGAACCGCCATGACAA 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
Qy 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGTTTAAACACACCCCGCCAGTTTGAAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
Qy 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTACTTCCAGGGCTTACGAGGAAATAG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
Qy 1010 TTGGAGGCGCCTCAATGCGGAATTCATGATGATACGAGGGTCTGTGTAGAGAGAT 1069

Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn	340
Qy	1070	GCAGAGGTGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGTCTCGAGTAGTGAGAAC	1129
Db	341	AlaGlyGlyGluGlnGluIleuArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
Qy	1130	AATCAAGAGAGTGATGATGCAAAAGTCTCAAAGGAGCAGCCTTGTAAGAACTTACTTAAGCACGCT	1189
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
Qy	1190	AAATCCGTCTCAAAAGAAAGGCTCCGAAGAAGAGGAGGATATCACCAACCCCAATCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
Qy	1250	AGACAAGCCGAGCCCGCATCTTTCTAAACAACCTTTGGGAAGTTATTGTAGGTGAAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
Qy	1310	AAGAAGAACCCCAAGCTTCAGGACCTCGACATGATGCTCACCTCTGTAGAGATCAAAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
Qy	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAAAGGCCATGGTTATTCGTCGTCGTCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	460
Qy	1430	GGAACCTGGAACCTTGAACCTCGTCGCTGTGAAGAAAGAGCAACAACAGAGGGGACGGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgLysArg	480
Qy	1490	GAAGAAGAGGAGGACGAAGACGAAGAAGAGGAGGAGGAAGTAACACAGAGGTGCGTAGGTAC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	500
Qy	1550	ACAGCAGGTTGAAGGAAGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
Qy	1610	AACGCTTCTCCGAACCTCCATCTCTCTGGCTTCGGTATCAACGCTGAAAAACAACACACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
Qy	1670	ATCTTCTCTTGCAAGTAAAGGACAAATGTCATAGACCAGATAGAGAAGCAAGCAAGCAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
Qy	1730	TTAGCATTCCTCGGTCCGGTGAAACAAGTTGAGAAGCTCATCAAAAAACCAAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
Qy	1790	CACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCAATCTCGTCTGCTCTCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
Qy	1850	TCTCCTCAGAAAGAGGATCAAGAGAGGAGAAACCAAGAGGAGGAGGTCCACCTCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluGlnGlnGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGGCTTTTAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 3			
ABUS2412			
ID ABUS2412 standard; protein; 626 AA.			
XX			
ABUS2412;			
XX			
XX			
DT 10-MAR-2003 (first entry)			
XX			
DE Peanut Ara h1 protein from cDNA clone P41b.			
XX			
KW Peanut; allergen; Ara h1; Ara h2; Ara h3; IgE binding site;			

KW	anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX	
OS	Arachis hypogaea.
XX	
PN	W0200274250-A2.
XX	
PD	26-SEP-2002.
XX	
PF	18-MAR-2002; 2002W0-US0009108.
XX	
PR	16-MAR-2001; 2001US-0276822P.
XX	
PR	18-MAR-2002; 2002US-00276822.
XX	
PA	(PANA-) PANACEA PHARM.
XX	
PI	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI	Rabjohn PA, Shin DS, Stanley JS;
XX	
DR	WPI; 2003-018765/01.
DR	N-PSDB; ABX70603.
XX	
PT	New modified anaphylactic food allergen, useful for preventing or
PT	treating allergic reactions associated with e.g. anaphylactic allergens.
XX	
PS	Claim 27; Fig 13; 300pp; English.
XX	
CC	The invention relates to a modified anaphylactic food allergen has an
CC	amino acid sequence that is substantially identical to that of natural
CC	anaphylactic food allergen, except for a cysteine residue that has been
CC	modified so that it cannot participate in the disulphide bond. The
CC	modification may also comprise mutation of the IGE binding sites to
CC	reduce allergenicity. Also included are: (1) a method of making a
CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC	or for causing a site specific mutation in the modified anaphylactic food
CC	allergen; (3) a transgenic plant or animal expressing the modified
CC	anaphylactic food allergen; (4) a method of treating an individual by
CC	reducing the clinical response to a natural anaphylactic food allergen.
CC	and an isolated fragment of peanut allergen Ara h 1. The modified
CC	anaphylactic food allergen is useful for preventing or treating allergic
CC	reactions associated with any natural allergen such as food, insect,
CC	rubber or preferably anaphylactic allergens. It is also useful for
CC	treating wounds in mammals such as bovine, canine, feline, caprine,
CC	ovine, porcine, murine or equine species. The present sequence is a
CC	peanut allergen (e.g. Ara h1, h2 or h3)
XX	
SQ	Sequence 626 AA;
Alignment Scores:	
Pred. No.:	1-4e-304 Length: 626
Score:	3286.00 Matches: 626
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	92.25% Indels: 0
DB:	6 Gaps: 0
US-10-728-323-1 (1-2032) x ABUS2412 (1-626)	
Qy	50 ATCAGAGGAGGGTTTTCACCTGATGCTGTGTGCTAGGATCCTTGCTCGTTCAGTT 10
Db	1 MetArgGlyArgValSerProLeuMetLeuLeuLeuValLeuAlaSerVal 20
Qy	110 TCTGCACCGNTGCCAAGTCATCCTTACAGAGAAACAGAAACCCCTGCCCCAG 16
Db	21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
Qy	170 AGGTGGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 22
Db	41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
Qy	230 CGTGCACCAAGCTCGAGTATGATCCTCGTTGTGCTATGATCTCGAGGACACACTGC 28
Db	61 ArgCysThrIysLeuGlnTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

```

QY 290 ACCACCAACCAAGTTTCCCTCCAGGGAGCGACACGTGGCCGCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCCCGCTCAACCCCAAGAGAGAGAGAGCGCATGGCGACCCAGCTGCACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCAACAGT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCAGGAATAAGGCCCGCAAGAGAGAGAGAGAGCAACAAGATGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CCAAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGAAACCAAAACCGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGCTCAAGGAGCTTTCGAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGCTTCCCAAGCAGCTGTAGCTGATAACATCTTGTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAAGAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCCATCCGGTTTCATTTCTCTACATCTTGAAACCGCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTACTTGCAGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATCGGAAATTCATAGATACGGAGGTGCTGTTAGAGAAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGAGAGAGAGCGAGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTAGTAGTCAAGTGTCAAGAGGACGACGTTCAAGAACTTACTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AATCCGCTTCAAGAAAGGCTCCGAAGAGAGAGGAGATATCAACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGGCGGCCCATCTTCTTAACAACCTTTGGAACTTATTTGAGGTGAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTCGACATCATCTACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

```

```

QY 1370 GGAGCTTTGATGCTCCCACTTCACTCAAAAGGCCATGTTATCGTCTCGTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAAACTTGAACCTCGTGGCTGTAAGAAAAGAGCAACAACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArg 480
QY 1490 GAAAGAGAGGAGGACGAAGACCAAGAGAGAGGAGGAAGTAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTCAAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACCGTTCCTCCGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGCAATGTGTAGTACCAGATAGAGAAAGCAAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCCGGTGCAACAGTTGAGAAGCTCATCAAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCGTCTGCTCTCAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAAAAACCAAGGAGGAGGTCCTCTCTTTC 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 4
ADC34906
ID ADC34906 standard; protein; 626 AA.
XX
AC ADC34906;
XX
XX 18-DEC-2003 (first entry)
DT
DE Peanut allergen Ara h 1.
KW allergen; antigen; hyporesponsive; desensitisation; immunomodulator;
KW gene therapy; peanut.
XX
OS Arachis hypogaea.
XX
XX WO2003047618-A2.
PN
XX 12-JUN-2003.
PD
XX 05-DEC-2002; 2002WO-CB005548.
PP
XX 05-DEC-2001; 2001US-0338385P.
PR
XX (CIRC-) CIRCASSIA LTD.
PA
XX Larche M, Ledger PW;
XX
XX WPI; 2003-523267/49.
DR
XX
PT Desensitizing an individual to a selected polypeptide antigen comprises
PT administering a composition containing polypeptide antigens in an amount
PT that generates a state of hyporesponsiveness to the antigen to allow
PT desensitization.

```


XX Disclosure; Page 40; 57pp; English.

PS The invention relates to a novel method for desensitising an individual

XX to a selected polypeptide antigen. The method comprises administering a

CC composition that contains polypeptide antigens in an amount that

CC generates in the individual a state of hyporesponsiveness to the antigen

CC to allow desensitisation to one or more polypeptide antigens. The method

CC of the invention has immunomodulator activity, and may have a use in gene

CC therapy. The composition and method are useful in manufacturing a

CC medicament for desensitising an individual to a selected polypeptide

CC antigen or for generating in the individual a state of hyporesponsiveness

CC to the antigen to allow desensitisation to one or more polypeptide

CC antigens. The present sequence is used in the exemplification of the

XX invention.

SQ Sequence 626 AA;

Alignment Scores:

Pred. No.:	1..4e-304	Length:	626
Score:	3286.00	Matches:	626
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.25%	Indels:	0
DB:	7	Gaps:	0

US-10-728-323-1 (1-2032) x ADC34906 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCAGTATGCTGTGTAGGATCCTTGTCTGCTCAGTT 109

DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20

QY 110 TCTGCAACCATGCGCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTGCCCCAG 169

DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGGTGCTCCAGATTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCGAGTCT 229

DB 41 ArgCysLeuGlnSerCysGlnGlnProArgAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTCGAGGACACACTGCG 289

DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGTTTCCCTCCAGGGAGCGGACACAGTGGCGGCCCAACCCCGAGCTAC 349

DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGGACCACTGGACCG 409

DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120

QY 410 AGGGAGCTGTAAGAGAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGGGCGACCAAGT 469

DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140

QY 470 CATCAGACCCACGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529

DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnIleArgGlyThr 160

QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCCCTTCTACTTCCCGTCAAGG 589

DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGTTTAGCACCGCTACGGGAACCAACCGTAGATCCGGGTCTCGAGAGTTTGAC 649

DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGGTCAAGCGAGTTTCCAGAAATCTCCAGAAATCAACCGTATTGTGCAGATCGAGGCCAAA 709

DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTCTTCCCAAGCAGCGCTGATGCTGTGAATCATCTCTTGTATCCAGCA 769

DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY 770 GGGCAAGCCACCGTACCGTAGCAATGCGCAATAACAGAAAGAGCTTTAACTTTCACGAG 829

DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTATCTTCTCATCTTCAACCGCATCACAAAC 889

DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTCAGGAT 949

DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCCTCCCGGCGAGCGGAGACCAATCATCTTCTTGCAGGGCTTCAGCAGGATACG 1009

DB 301 PhePheProAlaSerSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGAGAGCGGCTTCAATGCCGAATTCATAGATACGAGGGTCTCTTAAAGAGAGAAAT 1069

DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340

QY 1070 GCAGAGGTGAGCAAGAGAGAGAGGCGAGGCGATCGAGTACTCTCGAGTAGTGAGAAC 1129

DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360

QY 1130 AATGAGAGAGTATGATCARRAGTGTCAAGAGGAGGAGCGTGTGAAGACTTACTTAAGCACCT 1189

DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AAATCCGCTCAAGAAAGGCTCCGAGAGAGGAGAGATATCACCAACCCCAATCACTTG 1249

DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

QY 1250 AGAGAGCGAGCCCGATCTTTCTAAACACTTGGGAAGTTATTGTAGGTGAAGCCAGAC 1309

DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAAGACCCCGCTTCAGGACTCGACATGATGCTCCTCTGTAGAGACTCAAGAA 1369

DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCTGCTCGTCAACAAA 1429

DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460

QY 1430 GGAATCGAAACCTTGAACTCTGCTGCTATAAGAAAGCAACACAGAGGGGACGGCGG 1489

DB 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAGAGGAGGAG 1549

DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY 1550 ACAGCAGGTTGAAGAGAGCGATGTTCTATCATGCCAGAGCTCATCCAGTAGCCATC 1609

DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520

QY 1610 AACGCTTCTCCGAACTCCATCTCTGCTTGGCTTCCGATATCAACGCTGAAAAAACCAACA 1669

DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540

QY 1670 ATCTTCTTGCAGGTGATAAGGCAATGTGTATAGACCAAGAGAGAGAGAGAGAGAGAGAG 1729

DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560

QY 1730 TTAGCATTTCCCTGGGTCGGGTGAAACAGTTGAGAGGCTCATCAAAAAACAGAGGAATCT 1789

DB 561 LeuAlaPheProGlySerGlyGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGTCTGCTCTCTCGAGAGAGAG 1849

Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGAAACCAAGAGGAGGAGGTCACCTCTTCA 1909
 Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGGCTTTTAAAC 1927
 Db 621 IleLeuLysAlaPheAsn 626

RESULT 5

AAU04706
 ID AAU04706 standard; protein; 626 AA.

XX AAU04706;

DT 23-OCT-2001 (first entry)

XX Anaphylactic antigen Ara h 1.

XX Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse.

XX Mus sp.

EH Key Location/Qualifiers

FT Misc-difference 285 /note= "Encoded by gta"

XX W0200140264-A2.

XX 07-JUN-2001.

PF 06-DEC-2000; 2000MO-US033124.

XX 06-DEC-1999; 99US-00455294.

PR 23-JUN-2000; 2000US-0213765P.

PR 27-SEP-2000; 2000US-0235797P.

XX (PANA-) PANACEA PHARM LLC.

PA (UYAR-) UNIV ARKANSAS.

PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.

XX Bannan CA, Burks WA, Caplan MJ, Sampson H, Sosin H;

XX WPI; 2001-381378/40.

DR N-PSDB; AAS08537.

XX Antigenic fragments useful for reducing anaphylactic risk and reducing
 PT the severity and/or number of allergic symptoms in individuals sensitive
 PT to antigens, have reduced ability to bind Immunoglobulin E.

XX Claim 7; Fig 9; 100pp; English.

XX The sequence represents the amino acid sequence of anaphylactic antigen
 CC Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design
 CC antigenic peptides having a reduced ability to bind immunoglobulin E
 CC (iGE) as compared with the intact (A), or having a sequence substantially
 CC identical to a portion of sequence of an antigen that includes at least
 CC one iGE binding site, where at least one iGE binding site of the peptide
 CC is altered. The antigenic peptides are used in a composition which is
 CC useful for reducing risk or severity of allergic reaction to an antigen.
 CC This is done by identifying an individual at risk of allergic reaction to
 CC an antigen by identifying prior display of allergic symptoms when exposed
 CC to the antigen, or a familial relationship with an individual who
 CC previously displayed allergic symptoms when exposed to the antigen.
 CC Following this an antigen-specific iGE present on one or more mast cells
 CC or basophils in the individual's serum is identified. The individual is
 CC then contacted with a peptide corresponding to a portion of the antigen,
 CC which is selected, formulated, and delivered so that binding of the
 CC peptide to antigen-specific iGE is reduced as compared with iGE binding
 CC of intact antigen. The composition is also useful for treating and
 CC preventing allergic reactions

XX SQ Sequence 626 AA;

Alignment Scores:

Pred. No.: 3,38e-304 Length: 626
 Score: 3282.00 Matches: 625
 Percent Similarity: 99.84% Conservatives: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.14% Indels: 0
 DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x AAU04706 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCCTGTCTCTGGCTTCAGTT 109
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAAGCATGCCCAAGTCATCCTTACCAGAGAGAAACAGAGAACCCCTGCGCCAG 169
 Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTCCCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGCATGCGAGTCT 229
 Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
 Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGCTGGCGCCCAACCCCGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGAGAGGCGGATGGGGACCACTGGACG 409
 Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGGTGAAGAGAGAGAGACTGGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
 Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGACGCGCAAAATAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTTCTACTTCCCGTCAAG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGTTTAGACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCTGAGAGGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGCAGTTTCCAGATCTCCAGATCAACCGTATTGTGCAGATCCAGGCCAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTGTCTTCCCAAGCACCTGATGTGATAACATCTTGTATTCCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATTTGACGAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCACTCAGATCCCATCCGTTTCTTCTACATCTTCAACCGCATCAGCAAC 889
 Db 261 GlyHisAlaLeuArgGileProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTCAGGAT 949
 Db 281 GlnAsnLeuArgAlaAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

170 AGTGCCCTCCAGATTGCTCAACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db |||||
41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY CGTGTCAACCAAGCTCGAGTATGATCTCTGTTGTCTATCATCTCTCGAGACACACTGGC 289
Db |||||
61 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACAGTGGCGGCCCAACCCGAGACTAC 349
Db |||||
81 ThrAsnGlnArgSerProCglyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGCTCAACCCCGAAGAGAGAGAGGCGCGATGGCGGACCAAGTGGACCG 409
Db |||||
101 AspAspAspArgGlnProArgGlnGluGlyArgTyrProAlaGlyPro 120
QY 410 AGGAGCGCTCAAGAGAGAAGACTCGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT 469
Db |||||
121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
QY 470 CATCAGCAGCCCGGAAAATAAGCCCGAAGAGAGAGAGAGAACAAAGTGGGGAACA 529
Db |||||
141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGlnGlnTyrGlyThr 160
QY 530 CCAGTAGCCATGTAGGGGAGAACAATCTCGGAACAACCCCTTCTACTTCCCGTCAAG 589
Db |||||
161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCCCGCTACGGGAACCAACCGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db |||||
181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db |||||
201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTTCCAGCAGCAGCTGATGATAACAATCTCTGTTATCCAGCAA 769
Db |||||
221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCACAGCCACCTGACCGTAGCAATGCGCAATAGCAAGAGAGCTTTAATCTTTGACGAG 829
Db |||||
241 GlyGlnAlaThrValThrValAlaAsnGlyAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCATCAGAAATCCCATCCGTTTCAATTTCCATCTTGAACCGCCATGACAAC 889
Db |||||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGTAAATCTCCATCCCGCTTAAACACACCCCGCGCTTTGAGGAT 949
Db |||||
281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTCTCTGAGGCTTTCAGCAGGAATACG 1009
Db |||||
301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGAAATCAATGAGATACGGAGGCTGCTGTAGAGAGAAT 1069
Db |||||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGACGAGGAGAGAGGCGCAGAGCGATGGAGTACTCGGAGTAGTGAGAC 1129
Db |||||
341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGACGAGTGTAGTCAAGGTCTCAAGGACACGTTTCAAGAACTTACTTAGACGCT 1189
Db |||||
361 AsnGluGlyValIleValLysValSerLysGluArgGlyHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAGAAAGCTCCGAGAGAGGAGATATCACCAACCCCAATCAACTTG 1249
Db |||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

1250 AGAGAGGCGAGCCCGATCTTTCTAACACTTTGGGAAGCTTATTTGAGGTGAGCCAGAC 1309
Db |||||
401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGAGCTTCCAGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db |||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCTGCTGCTCAACAA 1429
Db |||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGACAACAACAGAGGCGCCGCG 1489
Db |||||
461 GlyThrGlyAsnLeuGluValAlaValArgLysGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGAGCAGAACGAGAGAGGAGGAAGTAACAGAGAGGTGCGCTAGGTAC 1549
Db |||||
481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTGAAGAGAGCGGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db |||||
501 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACACACACAGA 1669
Db |||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATTAAGGACAATGTGATAGACAGATAGAGAAAGCAAGGAGAT 1729
Db |||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGGTGAAAGTTGAGAGCTCATCAAAACCCAGAAAGGAATCT 1789
Db |||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTCAGTGTCTCTCAATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1849
Db |||||
581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db |||||
601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyGlyProLeuLeuSer 620
QY 1910 ATTTGAGCGCTTTTAAAC 1927
Db |||||
621 IleLeuLysAlaPheAsn 626
RESULT 7
ABU52568
ID ABU52568 standard; protein; 626 AA.
XX AC ABU52568;
XX DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h1 mutant K32A.
XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX OS Homo sapiens.
XX PN WO200274250-A2.
XX XX 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX XX

(PANA-) PANACEA PHARM.
 Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 Comedre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 Rabjohn PA, Shin DS, Stanley JS;
 WPI; 2003-018765/01.
 New modified anaphylactic food allergen, useful for preventing or
 treating allergic reactions associated with e.g. anaphylactic allergens.
 Example 5; Page; 300pp; English.
 The invention relates to a modified anaphylactic food allergen has an
 amino acid sequence that is substantially identical to that of natural
 anaphylactic food allergen, except for a cysteine residue that has been
 modified so that it cannot participate in the disulphide bond. The
 modification may also comprise mutation of the IGE binding sites to
 reduce allergenicity. Also included are: (1) a method of making a
 modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 or for causing a site specific mutation in the modified anaphylactic food
 allergen; (3) a transgenic plant or animal expressing the modified
 anaphylactic food allergen; (4) a method of treating an individual by
 reducing the clinical response to a natural anaphylactic food allergen;
 and an isolated fragment of peanut allergen Ara h 1. The modified
 anaphylactic food allergen is useful for preventing or treating allergic
 reactions associated with any natural allergen such as food, insect,
 rubber or preferably anaphylactic allergens. It is also useful for
 treating wounds in mammals such as bovine, canine, feline, caprine,
 ovine, porcine, murine or equine species. The present sequence is a food
 allergen, mutated to alter its IGE binding characteristics. Note: The
 present sequence is not shown in the specification but was created by the
 indexer using information provided in the specification
 Sequence 626 AA;
 Alignment Scores:
 Pred. No.: 5.25e-304 Length: 626
 Score: 3280.00 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.08% Indels: 0
 DB: 6 Gaps: 0
 US-10-728-323-1 (1-2032) x ABUS2568 (1-626)
 QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTGCTAGGATCTTGTCTCGCTTCAGTT 109
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAAGAGAAACAGAGAACCCCTGCCGCCAG 169
 Db 21 SerAlaThrHisAlaLysSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCCAGATGTTGTCACAGGAACCGGATGACTTGAAGCAAAAGCATGCCAGTCT 229
 Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCACTCGAGTATGATCTCGTGTGCTATGATCTCGAGGACACACTGGC 289
 Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCTCCCTCCAGGGAGCGGACACGTGGCGGCCCAACCCCGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGGAGCGCGATGGGGACCACTGGACCG 409
 Db 101 AspAspAspArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGGAGCGTGAAGAGAAAGCACTGGAGCAACCAAGAGAGATGGAGGCGACCAAGT 469
 Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140

QY 470 CATCAGCAGCCACGGAATAAAGGCCGGAAGAGAGAGAGAGAAACAAGAGTGGGAACA 529
 Db 141 HisGlnGlnProArgGlyLeileArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTACACCCGCTACGGGNAACCAAAACGGTAGGATCCGGTCTCCTCAGAGGTTTCAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCAACCGTATTGTGCAGATCCAGGCCAAA 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTGTTCCTCCAAAGCACGCTGATGATACATCCTTGTATTCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTCAGCAGTAAATCGCAATACAGAAAGAGCTTTAATCTTCACGAG 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTATCTTCTATACATCTTGAACCGCATCACAC 889
 Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACTCAGAGTAGCTAAATCTCATGCGGTTAACACACCCGCGCAGTTTGAGAT 949
 Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTTCTTCCCGGCGAGCAGCCGAGACCAATCATCTTCTAGGAGGCTTCAGAGGAGTAACG 1009
 Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGAGAGCGCGCTTCAATGCGGAATTCATAGATACACGAGGAGTCTGTAGAGAGAGAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGAGGTGAGCAAGAGAGAGAGGAGGAGGAGGAGGAGTACTCGAGTAGTAGAGAAC 1129
 Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGCAGGTATAGTCAAAAGTGTCAAAGGAGACGTTGAAGAACTTACTAAGCACGCT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGTCTCAAAGAAAGGCTCCGAAAGAGGAGAGATATCACCAACCCCAATCAACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAGGCGAGCCCGATCTTCTAAACACTTTGGGAAGTTATTGTAGGTGAAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLysSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAGAACCCCGCTCAGACCTGACATGATGCTCACTGTGTAGAGATCAAGAA 1369
 Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLysGlu 440
 QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCTCGTCAACAAA 1429
 Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
 QY 1430 GGAAGTGAAGACCTTGAATCTCGTGGCTGTAGAAAGAGCAACACAGAGGCGCGGG 1489
 Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
 QY 1490 GAAGAGAGGAGGAGCAAGAGCAAGAGAGGAGGAGGAGTAAACAGAGAGTGGTGGTAC 1549
 Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY	770	GGCAAGCCACCGTAGCCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAenGlyAenAenArgLysSerPheAenLeuaspGlu	260
QY	830	GGCCATGCACTCAGATCCATCCGGTTTCATTCCTTACATCTTGAAACCGCCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAenArgHisaspAen	280
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCGTAAACACACCCGCCAGTTTCAGGAT	949
Db	281	GlnAenLeuArgValAlaLysIleSerMetProValAenThrProGlyGlnPheGluasp	300
QY	950	TTCTTCCCGCCAGCAGCCGACCAATCATCTTCTCAGGGCTTCAGCAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgaspGlnSerSerTyrLeuGlnGlyPheSerArgAenThr	320
QY	1010	TTGGAGCGCCCTTCAATCGGAATTCATAGATACGGAGGGTGTGTTGAAAGAGAT	1069
Db	321	LeuGluAlaAlaPheAenAlaGluPheAenGluIleArgArgValLeuLeuGluAen	340
QY	1070	GCAGGAGGTGACCAAGAGCAGAGCGGACGAGCGGATGAGTACTCGGAGTAGTGAGAC	1129
Db	341	AlaGlyGlyGluGlnGluArgGlyGlnArgTyrSerThrArgSerSerGluAen	360
QY	1130	AATGAAGGAGTGATAGTCAAAAGTGTCAAAAGGAGCAGCTTTGAAGAACTTACTAAGCACGCT	1189
Db	361	AenGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAAAGAAAGCCGCGAAGAGGGAGATATCACCAACCCCAATCAACTTG	1249
Db	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAenProIleAenLeu	400
QY	1250	AGAGAGGCGACCCGATCTTTCTAACACTTTGGGAAGTTATTTGAGGTGAGCCAGAC	1309
Db	401	ArgGluGlyGluProAspLysSerAenAenPheGlyLysLeuPheGluValLysProasp	420
QY	1310	AAGAAGAACCCCGAGCTTCAGAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAenProGlnLeuGlnAspLeuAenMetMetLeuThrCysValGluLysGlu	440
QY	1370	GGAGCTTTGATGCTCCCACTTCAACTCAAGGCCATGGTTATCGTCTGCTCAACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAenSerLysAlaMetValIleValValAenLys	460
QY	1430	GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAGACCAACAGAGGGGACGGCG	1489
Db	461	GlyThrGlyAenLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAGAGGAGGACGAAGACGAAGAGGAGGAAAGTAAACAGAGAGTGCGTAGGTAC	1549
Db	481	GluGluGluGluaspGluaspGluGluGluGluGlySerAenArgGluValArgArgTyr	500
QY	1550	ACAGCAGGTTGAAGAACGGATGTTTCATCATGCGCAGCAGCTCATCCATGAGCCATC	1609
Db	501	ThrAlaAenLeuLysGluGlyaspValPheIleMetProAlaAHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTGGTATCAAGCTCAAAACACACACAGA	1669
Db	521	AenAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAenAlaGluAenAenHisArg	540
QY	1670	ATCTTCTTGCAGGTGATAGGACAAATGTGATAGACCCAGATAGAGAAGCAAGCGAAGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAenValIleaspGlnIleGluLysGlnAlaLysasp	560
QY	1730	TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAACACAGAAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAenGlnLysGluSer	580
QY	1790	CACCTTGTGAGTGCCTCCTCAATCTCAATCTCAATCTCCGTCGTCCTCCAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600

QY	1850	TCTCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGTCACTCCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGlnGlyGlyGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAC	1927
Db	621	IleLeuLysAlaPheAen	626
RESULT 9			
ABUS574	standard; protein; 626 AA.		
XX	ABUS2574;		
DT	10-MAR-2003	(first entry)	
XX	Peanut Ara h1 mutant R499A.		
DE	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;		
KW	muten; anaphylactic food allergen; antiallergenic; vaccine;		
KW	wound healing.		
XX	Homo sapiens.		
OS	WO200274250-A2.		
PN	26-SEP-2002.		
PD	18-MAR-2002; 2002WO-US0009108.		
XX	16-MAR-2001; 2001US-0276822P.		
PR	18-MAR-2002; 2002US-00276822.		
XX	(PANA-) PANACEA PHARM.		
PA	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;		
XX	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
PI	Rabjohn PA, Shin DS, Stanley JS;		
XX	WPI; 2003-018765/01.		
DR	New modified anaphylactic food allergen, useful for preventing or		
XX	treating allergic reactions associated with e.g. anaphylactic allergens.		
PT	Example 5; Page; 300pp; English.		
PS	The invention relates to a modified anaphylactic food allergen has an		
XX	amino acid sequence that is substantially identical to that of natural		
CC	anaphylactic food allergen, except for a cysteine residue that has been		
CC	modified so that it cannot participate in the disulphide bond. The		
CC	modification may also comprise mutation of the IgE binding sites to		
CC	reduce allergenicity. Also included are: (1) a method of making a		
CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding		
CC	or for causing a site specific mutation in the modified anaphylactic food		
CC	allergen; (3) a transgenic plant or animal expressing the modified		
CC	anaphylactic food allergen; (4) a method of treating an individual by		
CC	reducing the clinical response to a natural anaphylactic food allergen;		
CC	and an isolated fragment of peanut allergen Ara h 1. The modified		
CC	anaphylactic food allergen is useful for preventing or treating allergic		
CC	reactions associated with any natural allergen such as food, insect,		
CC	rubber or preferably anaphylactic allergens. It is also useful for		
CC	treating wounds in mammals such as bovine, canine, feline, caprine,		
CC	ovine, porcine, murine or equine species. The present sequence is a food		
CC	allergen, mutated to alter its IgE binding characteristics. Note: The		
CC	present sequence is not shown in the specification but was created by the		
CC	indexer using information provided in the specification		
XX	Sequence 626 AA;		
SQ			
Alignment Scores:			
Pred. No.:	5.25e-304	Length:	626
Score:	3280.00	Matches:	625
Percent Similarity:	99.84%	Conservative:	0

Best Local Similarity: 99.84%		Matches: 1	
Query Match: 92.08%		Indels: 0	
DB: 6		Gaps: 0	
US-10-728-323-1 (1-2032) x ABUS2574 (1-626)			
QY	50	ATCAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCCTTGTCTCGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGATGCCAAGTCATCACCTTACCAAGAGAAACAGAGAACCCCTCGCCACG	169
DB	21	SerAlaThrHisAlaLysSerProTyrGlnLysIleThrGluAsnProCysAlaGln	40
QY	170	AGTGCTCTCAGAGTTGTCAACAGGAACCGATGACTTGAAGCAAAAGGATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCAGATATGATCTCTGTTGTCTATGATCTTCGAGGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACAGCTGGCCGCCAACCCGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATCACCGCTCAACCCCAAGAGAGAGAGGCGCATGGGGAACAGCTGGACCG	409
DB	101	AspAspAspArgGlnProArgGluGluGlyGlyArgTyrGlyProAlaGlyPro	120
QY	410	AGGAGCGTGAAGAGAGAGAGCTGGAGACCAACCAAGAGAGATTTGAGGCGGACCAAGT	469
DB	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	140
QY	470	CATCAGCAGCCAGGAATATAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr	160
QY	530	CCAGGTAGCATGTGAGGAGAGAAACATCTCGGAACAAACCTTCTACTTCCCTCGAAG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCGCTACGGGAACCAACCGTAGGATCCGGTCTCGCAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCAAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCAAGCACCGTACCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCATGCACTCAGATCCATCCCGTTTCATTTCTATCATCTTGAACCCGCTATGACAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCCCGTTTAAACACACCCGGCCAGTTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGCAGCCGAGACCAATCATCTCTACTTGCAGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGGCGCCCTTCAATCGGAATTCATGATACGAGAGAGGCTGCTTGAAGAGAAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
QY	1070	GCAGGAGGTGACCAAG	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
QY	1130	AATGAAGAGTGTATAGTCAAAAGTGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAGAGAGAGAGAGATATCAACCAACCAATCACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGAGCCCGATCTTTTAAACAACCTTTCGGAAAGTTATTGTAGGTCAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAGAACCCCGCTCCAGACCTCGACATGATGCTCACCTGTGTGAGAGATCAAAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCCACTTCAACTCAAAAGGCCATGGTTATCTGCTGCTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAATCGAAACCTTGAATCTGCTGCTGTGAAGAAAGAGCAACAACAGAGGGAGCGCGG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAGCGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGlu	500
QY	1550	ACAGCGAGGTTGAAGAGAGGCGATGTTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTCTGCTTGGCTTCCGTTATCAACGCTGAAAAACAACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTCTGAGGTGATAAGGACAAATGTGTATAGACCAAGATAGAGAGAGCAAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTTCCCTGGTTCGGGTGAACCAAGTTGAGAAAGCTCATCAAAACAGAGGAATCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCCTCTCGAGAAAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlyLysGlu	600
QY	1850	TCTCTCTGAGAAAGAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1909
DB	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC 1927	
DB	621	IleLeuLysAlaPheAsn 626	
RESULT 10			
ABUS2573			
ID	ABUS2573	standard; protein; 626 AA.	
XX	XX	ABUS2573;	
AC	XX		
DT	10-MAR-2003	(first entry)	
XX	XX	Peanut Ara h1 mutant R109A.	
DE	XX		
XX	XX	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;	
KW	XX	mucin; anaphylactic food allergen; anti-allergenic; vaccine;	

KW wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO200274250-A2.
 XX
 PD 26-SEP-2002.
 XX
 XX 18-MAR-2002; 2002WO-US009108.
 XX
 PF 16-MAR-2001; 2001US-0276822P.
 PR 18-MAR-2002; 2002US-00276822.
 XX
 XX (PANA-) PANACEA PHARM.
 PA
 XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compdre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabbjohn PA, Shin DS, Stanley JS;
 XX
 DR WPI; 2003-018765/01.
 XX
 XX New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 PT
 XX Example 5; Page; 300pp; English.
 XX
 XX The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IGE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a food
 CC allergen, mutated to alter its IGE binding characteristics. Note: The
 CC present sequence is not shown in the specification but was created by the
 CC indexer using information provided in the specification
 XX
 SQ Sequence 626 AA;

 Alignment Scores:
 Pred. No.: 5-25e-304 Length: 626
 Score: 3280.00 Matches: 625
 Percent Similarity: 99.84% Conservatives: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.08% Indels: 0
 DB: 6 Gaps: 0

 US-10-728-323-1 (1-2032) x ABUS2573 (1-626)

 QY 50 ATGAGAGGAGGTTTCTCCATGATGCTGTGCTAGGATCCTGTCTGCTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20

 QY 110 TCTGCAACCATGCCAAGTCATCCTTACCAAGAGAGAAACAGAGAACCCCTGCCCCAG 169
 DB 21 SerAlaIleHisAlaLysSerSerProTyrGlnLysLysLysLysLysLysLysLysLys 40

 QY 170 AGGTGCTCCAGATTGTCAACAGGAACCGGATGACTTCAAGCAAAAGGCATCCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60

 QY 230 CGCTGCACCAAGTCAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGC 289
 DB

Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGAGGACACGTCGGCCGCAACCCCGAGACTAC 349
 Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 409
 Db 101 AspAspAspArgGlnProArgAlaGluGluGlyArgGlyArgGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAGAAGAGAGAGCTGGAGACAACCAAGAGAGAGATTGGAGGCGACCAAGT 469
 Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGCAGCCAGCAAAATAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnIleTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGAGAGAGAAACATCTCGAACAACCCCTTCTACTTCCCGTCAAGG 589
 Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGCTAGATCCGGGTCTCGCAGAGGTTTGAC 649
 Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGGTCAAGCAGCTTTCAGAACTCCAGAACTCCAGAACTCCAGAACTCCAGAACTCCAG 709
 Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CTTAACTCTTGTCTTCTCCCAAGCAGCTGATGCTGATAACATCCTTGTATTCAGCAA 769
 Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCAGCTGACCGTACGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCA 829
 Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GCGCATGCACTCAGAACTCCATCCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 889
 Db 261 GlyHisAlaLeuArgGlyProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAAACACACCCGCGCAGTTTTCAGGAT 949
 Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTTCTTCCCGGAGCAGCGCAGACCAATCCTACTTCTGAGGCGCTTCTGAGGAGGATACG 1009
 Db 301 PhePheProAlaSerSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGCGCGCTTCAATGCGGAATTCATGAGATACGAGGAGTCTGTAGAGAGAT 1069
 Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGGAGTGAGCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1129
 Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGGAGTGATGATCAAGTGTCAAAGGAGACGCTTGAAGAACTTACTTAAGCACCGT 1189
 Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AAATCCGCTCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCAACCAATCAACTTG 1249
 Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAGAGCGAGCCCGATCTTTCTTAACAACTTTGGAAAGTTATTTCAGGTGAAGCCAGAC 1309
 Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAGACCCCGAGCTTCAGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
 Db 421 LysLysAsnProGlnLeuAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACACTTCAACTCAAGGCCATGGTTATCTCGTCTCGTCAACAAA 1429
 Db |||||
 QY 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
 Db |||||
 QY 1430 GGAACCTGGAACCTTGAACCTGCTGCTGTAAGAAAGAGCAACAAACAGAGGGGACGGCGG 1489
 Db |||||
 QY 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgArgArg 480
 QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
 Db |||||
 QY 481 GluGluGluGluAspGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 500
 QY 1550 ACAGCGAGGTGAGGAGGCGATGTTCTTCATCATGCGCAGCAGCTCATCCAGTAGCCATC 1609
 Db |||||
 QY 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGGAACCAACCAACAGAGA 1669
 Db |||||
 QY 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCTTCAGGTGATAGGACAAATGTGATAGCAGATAGAGAAAGCAAGCGAAGGAT 1729
 Db |||||
 QY 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCCTGGTGGTGACAAAGTTGAGAGCTCATCAAAACCAAGAGGATCT 1789
 Db |||||
 QY 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTGAAGAAGAG 1849
 Db |||||
 QY 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGGAGGGTCCACTCTTTCA 1909
 Db |||||
 QY 601 SerProGluLysGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 620
 QY 1910 ATTTGAGGCTTTTAAAC 1927
 Db |||||
 QY 621 IleLeuLysAlaPheAsn 626
 RESULT 11
 ADM12053
 ID ADM12053 standard; protein; 626 AA.
 AC AC
 XX ADM12053;
 XX
 XX 20-MAY-2004 (first entry)
 XX
 DE Arachis hypogaea 1 (Ara hi) protein.
 XX
 KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
 KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
 KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
 KW antiinflammatory; antidiabetic; antithyroid; immune;
 KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
 KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
 KW scleroderma; dermatomyositis; pemphigus.
 XX
 OS Arachis hypogaea.
 XX
 XX US2003235594-A1.
 PN
 XX 25-DEC-2003.
 PD
 XX 17-SEP-2002; 2002US-00245871.
 XX
 XX 14-SEP-1999; 99US-00396813.
 PR
 XX 17-JUL-2002; 2002US-00197000.
 XX
 XX (ANTI-) ANTIGEN EXPRESS INC.
 PA
 XX Humphreys R, Xu M;
 PI

XX WPI; 2004-070554/07.
 DR
 XX Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
 PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
 PT erythematosus and diabetes mellitus.
 XX
 XX Example 1; Page 16-17; 87pp; English.
 PS
 XX The invention relates to a novel antigen presentation enhancing hybrid
 CC polypeptide. The novel polypeptide has an N-terminal element consisting
 CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
 CC deletion modifications, a chemical structure covalently linking the N-
 CC terminal element to an MHC class II-presented epitope of a C-terminal
 CC element. The C-terminal element comprises an antigenic epitope, which
 CC binds to an antigenic peptide binding site of an MHC class II molecule.
 CC The antigen presentation enhancing hybrid polypeptide has the following
 CC activities: antibacterial, virucide, fungicide, antirheumatic,
 CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
 CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
 CC enhancing hybrid polypeptide is useful for modulating the immune response
 CC in an individual and for treating infections (such as bacteria, virus,
 CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
 CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 CC represents a mammalian Ii key related protein of the invention.
 XX
 SQ Sequence 626 AA;

Alignment Scores:
 Pred. No.: 6,54e-304 Length: 626
 Score: 3279.00 Matches: 625
 Percent Similarity: 99.84% Conservative: 0
 Best Local Similarity: 99.84% Mismatches: 1
 Query Match: 92.06% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-1 (1-2032) x ADM12053 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTCTCTGGTTCAGTT 109
 Db |||||
 QY 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGCATGCCAAGTCATCCTTACCAGAGAAACAGAGAACCCCTCGGCCAG 169
 Db |||||
 QY 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
 Db |||||
 QY 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlyLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTCGAGACACACTGGC 289
 Db |||||
 QY 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACGTGGCGCCGCAACCCGAGACTAC 349
 Db |||||
 QY 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGCTCAACCCGAGAGAGGAGGAGGCGCGGAGCCAGCTGGACCG 409
 Db |||||
 QY 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAGAAGAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGCAAGT 469
 Db |||||
 QY 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
 QY 470 CATCAGCAGCCACGGAATAAGCCCCGAGAGAGAGAGAGAGAAACAGAGTGGGGAACA 529
 Db |||||
 QY 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluGlnGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTCTTACTTCCGTCAGG 589
 Db |||||

161	ProGlySerHisValArgGluGluThrSerArgAsnAsnPropheTyrPheProSerArg	180
590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCCTCGACAGGTTTGAC	649
181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
650	CAAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAATCACCGTATTTGTGCGAGATCGAGGCCAAA	709
201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
710	CCTAACACTCTGTCTTCCCAAGCAGCTGATCCTGTATACATCTCTGTATTATCCACGAA	769
221	ProAsnThrLeuValLeuProGlyHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
770	GGGCAAGCCACCGTCGACGTCAGCAAAATGGCAATACACAAAGAGCTTTAATCTTTGACGAG	829
241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
830	GGCCATGCACTCAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCCGCATGACAAC	888
261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
890	CAGAACCTCAGAGTAGCTAAATCTCCATGCGCCCTTAACACACCGCGCCAGTTTGAGGAT	949
281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
950	TTCTTCCGGCGAGCAGCGCAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATACG	1009
301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
1010	TTGAGGCGCGCTTCAATGCGGAATCAATGAGATACGGAGGTCGTCTTAGAAGAGAAT	1069
321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
1070	GCAGGAGTGACGAAGAGGAGAGGCGGATGCGAGTCTCGAGTAGTGAGAAC	1129
341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
1130	AATCAAGGAGTGATAGTCAAAAGTGTCAAAAGGAGCAGCTTGAGAACTTACTTAAGCAGCT	1189
361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
1190	AAATCCGCTCTCAAAAGAAAGGCTCCGAAGAGAGAGGAGATATCACCAACCCCAATCAACTTG	1249
381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
1250	AGAAAGCGGAGCCCGATCTTTCTAAACAACTTTCGGGAAGTTATTGTAGGTGAGCCAGAC	1309
401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
1310	AAGAAGAACCCCAAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
1370	GGAGCTTTGATGCTCCACACTTCAAGGCCATGGTTATTCGTCTGTCGTCAACAAA	1429
441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	460
1430	GGAACTGGAACCTTTGAACCTCGTCGGCTGTAAAGAAAGAGCACAACAGAGGGGACGGCGG	1489
461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
1490	GAACAAGGAGGACGAGACGAGAGACGAGGGAGAGTAAACAGAGGTCGTCAGTAC	1549
481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	500
1550	ACAGCAGGTTGAAGGAAGGCGATGTGTTCATCATGTCAGCAGCTCATCAGTAGCCATC	1609
501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
1610	AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACCAAGA	1669
521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540

Qy	1670	ATCTTCCTTCAGGTGATAAGGCAATGTGATAGACAGATAGAGAAAGCAAGCAAGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
Qy	1730	TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAGCTCATCAAAAACCAAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValIleGluLeuIleLysAsnGlnLysGluSer	580
Qy	1790	CACTTTCGTGAGTGTCTGCTCTCAATCTCAATCTCCGTCGTCCTCTGAGAAAAGAG	1849
Db	581	HisPheValSerAlaIleArgProGlnSerGlnSerProSerSerProGluLysGlu	600
Qy	1850	TCTCCTCAGAAAAGAGTCAAGAGGAGGAGAAAACCAAGGAGGAGGTCACACTCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
Qy	1910	ATTTTGAAGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 12			
ABU52569	ABU52569 standard; protein; 626 AA.		
XX	AC	ABU52569;	
XX	AC		
DT	10-MAR-2003	(first entry)	
DE	Peanut Ara h1 mutant D52A.		
KW	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;		
KW	mutin; anaphylactic food allergen; antiallergenic; vaccine;		
KW	wound healing.		
XX			
OS	Homo sapiens.		
XX			
FN	W0200274250-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	18-MAR-2002; 2002WO-US009108.		
XX			
PR	16-MAR-2001; 2001US-0276822P.		
PR	18-MAR-2002; 2002US-00276822.		
XX			
PA	(PANA-) PANACEA PHARM.		
XX			
PI	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;		
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
PI	Rabjohn PA, Shin DS, Stanley JS;		
XX			
DR	WPI; 2003-018765/01.		
XX			
PT	New modified anaphylactic food allergen, useful for preventing or		
PT	treating allergic reactions associated with e.g. anaphylactic allergens.		
XX			
PS	Example 5; Page; 300pp; English.		
XX			
CC	The invention relates to a modified anaphylactic food allergen has an		
CC	amino acid sequence that is substantially identical to that of natural		
CC	anaphylactic food allergen, except for a cysteine residue that has been		
CC	modified so that it cannot participate in the disulphide bond. The		
CC	modification may also comprise mutation of the IgE binding sites to		
CC	reduce allergenicity. Also included are: (1) a method of making a		
CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding		
CC	or for causing a site specific mutation in the modified anaphylactic food		
CC	allergen; (3) a transgenic plant or animal expressing the modified		
CC	anaphylactic food allergen; (4) a method of treating an individual by		
CC	reducing the clinical response to a natural anaphylactic food allergen;		
CC	and an isolated fragment of peanut allergen Ara h 1. The modified		
CC	anaphylactic food allergen is useful for preventing or treating allergic		
CC	reactions associated with any natural allergen such as food, insect,		

CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification.
XX
SQ Sequence 626 AA;

Alignment Scores:

Pred. No.: 8,15e-304 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.84% Conservative: 0
Best Local Similarity: 99.84% Mismatches: 1
Query Match: 92.03% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABUS2569 (1-626)

QY	50	ATCAGAGGGAGGGTTCTCCACGATGCTGTTCTAGGATCCTTGTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal	20
QY	110	TCGCAACGCATCCCAAGTCATCACCTTACCAGAAGAAACACAGAACCCCTCGCCCCAG	169
DB	21	SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln	40
QY	170	AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGCAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer	60
QY	230	CGTGCAACGAGTCGAGTATGATCTCTGTTGTCTATGATCCTCGAGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCCTCCAGGGAGCGACACGTCGCGCCCAACCCGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluLysGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCGCGTCAACCCCGAAGAGAGGAGGCGGATGGGGACAGCTGGACCG	409
DB	101	AspAspArgArgGlnProArgArgGluGlyGlyArgTyrProAlaGlyPro	120
QY	410	AGGACGCTGAAGAGAGAGACTCGAGACCAACAGAGAGATGGAGGCGACCAAGT	469
DB	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer	140
QY	470	CATCAGCAGCCAGGAAATAAGCCCGAAGGAAGAGAGGAGAAACAAAGATGGGGAACA	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGlnGlnTyrGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCGCTACGGAACCAAAACCGTAGGATCCGGCTCTCGAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAGCTCAAGCAGTTTCAGATCTCCAGATCCACGCTATCTGTCAGATCCGAGCCAAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTCTTCCCAAGCACGCTGATGCTGATAACATCTCTTATTCAGACAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCAAGCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGCTTTTAACTTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCATCAGATCCCATCCGTTTCTATCTTACATCTTGACCCCATGACAC	889

DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACTTCAGAGTAGCTAAAAATCTCCATGCCCCGTAAACACACCCCGCCAGTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGCGAGACCAATCATCTTCTTCCAGGGCTTCAGCAGGATAGC	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGGAGGCGCCCTTCAATCGGAAATTCATAGATACGAGGGTGTCTGTAGAGAGAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340
QY	1070	GCAGAGGTGACAAAGAGAGAGAGCGGATGAGTACTCGGAGTGTAGTAGAAC	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn	360
QY	1130	AATGAGGAGTATAGTCAAACTGTCNAAGGAGCAGTTGNAAGACTTACTAAGCAGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTAAAGAAAGGCTCCGAAGAGAGGAGATATCAACCAACCAATCAACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGAGCCCGCATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATCTCCACACTTCAACTCAAAAGGCCATGTTATCTGCTCGTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	460
QY	1430	GGAACTGGAAACCTTGAACCTCGTGTCTGTAAGAAAGAGCAACAACAGAGGGAGCGCG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAGAGGAGGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCAGGTTTGAAGAAAGCGATGTTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACCTCATCTGCTTGGCTTATCAACGCTGAAACCAACCAACA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTTCAGAGTGATAAGGACATGTGTAGTACACAGATAGAGAGAGAGAGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTTCCCTGGGTGCGGTGAACAGTTGAGAGCTCATCAAAACCAAGAGGATCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTTGTGAGTGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGAGAAAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu	600
QY	1850	TCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGAGGAGGAGGAGGAGG	1909
DB	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC	1927
DB	621	IleLeuLysAlaPheAsn	626

QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGAGCGCCGATGGGGACCACTGGACCG 409
DB 101 AspAspAspArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGACGCTGAAGAGAACGACTGGACACCAACCAAGAGAGATGGAGGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCACCCACGGAAAAATAAGGCCGGAAGGAAGAGAGAGCAACAGAGTGGGGAAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CGAGTAGCCATGTGAGGGAAGAAACATCTCGGAACACCCCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTTGTTCTCCCAAGCACGCTGATGTGATGAATCAATCTTGTATTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAGCCACCGTAGCCCTAGCAATGCGCAATACAGAAAGAGCTTTAATCTTACCGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCCATCCGCTTTCATTTCTACATCTTGAAACCGCCATCACAAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCGCGTTAACACACCGCGCCAGTTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValThrGlnProGlyGlnPheGluAsp 300
QY 950 TTTCTTCCCGCGAGCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCGCTTCAATCCGAATTCATAGATACGGAGGTGCTGTGAAGAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGCGGACAGCGGATGAGTACTCGGAGTGTAGTAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTGTATGATCAAGTGTCAAGGAGCAGCTTGAAGACTTACTTAAGCACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGCGGACCCGATCTTTCTAACACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAACCCCGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGGCCATGGTTATCGTCTCGTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACTGGAAA CTTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGG 1489

DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 CAAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGGAAGTAAACAGAGAGTGGTGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTTGAAGAAAGCGCATGTGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTTCTCCGAACCTCATCTGCTTGGCTTCGTTATCAAGCTGAAACACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAAGCAACATGTGTAGACCAGATAGAGAAGCAAGCGAAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGGTCCGGTGAAACAAAGTTGAGAAGCTCATCAAAACCAAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCGTCTGCTCTCTGAGAAAAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909
DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
DB 621 IleLeuLysAlaPheAsn 626
RESULT 15
ADG27464
ID ADG27464 standard; protein; 625 AA.
XX
AC ADG27464;
XX
DT 26-FEB-2004 (first entry)
DE Peanut Ara h1 cDNA clone P41b protein.
XX
KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
KW antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.
XX
OS Arachis hypogaea.
XX
PN US2003202980-A1.
XX
PD 30-OCT-2003.
XX
PF 18-MAR-2002; 2002US-00100303.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-0071793P.
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.
PR 13-FEB-1998; 98US-0074633P.
PR 29-JUN-1998; 98US-00106872.
PR 27-AUG-1998; 98US-00141220.
PR 13-NOV-1998; 98US-00191593.
PR 29-JAN-1999; 99US-00240557.
PR 29-JAN-1999; 99US-00241101.
PR 11-FEB-1999; 99US-00248673.
PR 02-MAR-1999; 99US-00248674.
PR 02-MAR-1999; 99US-0122450P.
PR 02-MAR-1999; 99US-0122452P.
PR 02-MAR-1999; 99US-0122560P.
PR 02-MAR-1999; 99US-0122565P.

Db 361 AsnGluGlyValIleValIysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAGAAAGCTCCGAAGAGGAGATATCACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCGATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACGTGAAACCTTGAACTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGGGACGGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGAGGACCAAGACGAGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGCGGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACCAACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTGCAGCTCATPAGGACATGTGATAGACCATAGACAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGly---LysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 559
QY 1730 TTAGCATTCCTCGGTGGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAGAAAGGAATCT 1789
Db 560 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 579
QY 1790 CACTTTGTGAGTCTGCTCCTCAATCTCAATCTCAATCTCCGTGCTCTCTGAGAAAGAG 1849
Db 580 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 599
QY 1850 TCTCCTGAGAAAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 600 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 619
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 620 IleLeuLysAlaPheAsn 625

Search completed: August 24, 2005, 09:55:54
Job time : 154.082 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	1282	36.0	605	2	S20007	beta-conglycinin a
2	1281	36.0	605	1	FWGYBA	beta-conglycinin a
3	1267	35.6	571	2	S00566	convicillin precurs
4	1251	35.1	639	2	B24810	beta-conglycinin a
5	1178.5	33.1	459	2	S08505	vicillin - garden p
6	1174.5	33.0	433	2	S00567	vicillin precursor
7	1156	32.5	439	1	FWSYCB	beta-conglycinin b
8	1148	32.2	438	2	S35757	vicillin, 47K - gar
9	1109.5	31.1	463	2	A27288	vicillin precursor
10	1108.5	31.1	463	2	S08309	vicillin precursor
11	1068	30.0	410	1	FWPMVB	vicillin B precurs
12	962.5	27.0	445	2	JQ2264	canavalin - jack b
13	959.5	26.9	445	2	S00281	canavalin - sword
14	907.5	25.5	410	2	T44430	protein pV100 [imp

QY 230 CGGTGACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
||||| : : : : :
Db 57 ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyGluIleProArg--- 75
QY 263 GTCTATGATCTCGAGGACACACTGGACACCACCAACAGTTCCCTCAGGGGAGCG 322
||||| : : : : :
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACAGTGGCGCCAAACCGGAGACTACGATGAC----- 358
: : : : : : : : : : :
Db 85 ArgGluProGlnGlnProGlyGluLysGluLysGluLysGluLysGluLysGluLys 104
QY 359 -----CGCGTCAACCCCGAAGAGAG-----GAAGAGAGCCGA 391
||| : : : : : : : : : : :
Db 105 ProPheProArgProGlnProArgGlnGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACACGCTGGACCGAGGAGCGTGAAGAGAGAGAA 430
||| : : : : : : : : : : :
Db 125 TrpProArgLysGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144
QY 431 GACTGGAGACAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCATCAGCAGCCACGG 484
||| : : : : : : : : : : :
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProPheGlnLysGluGlu 164
QY 485 AAAATAAGGCCGAAGAGAGAGAGAACAGAGTGGGGAAACACAGGTAGCCATGTG 544
: : : : : : : : : : :
Db 165 ArgLysGlnGluGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGAA-----GAAACATCTCGGAACAACCCCTTTCTATTCCTCCGTCAGGCGGTTTAGC 598
||| : : : : : : : : : : :
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheLysSerAsnArgPheGlu 204
QY 599 ACCCGCTACGGGAACCAACGCTAGATCCGGTCTCGACAGGTTTCACCAAGGTCA 658
||| : : : : : : : : : : :
Db 205 ThrLeuPheLysAsnGlnTyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGCAGTTTCAGAAATCTCCAGATACCCGTATTGTGCAGATCGAGCGCAAACTCAACT 718
||| : : : : : : : : : : :
Db 225 ProGlnLeuGlnAsnLeuArgAspTyArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTGTGTTCTCCCAAGCAGCTGATGTATACATCTCTGTTATCCAGCAAGGCGAAGCC 778
||| : : : : : : : : : : :
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTGACCGTAGCAATGCATACAGAAAGAGCTTAACTTTCAGCAGGSCCATGCA 838
||| : : : : : : : : : : :
Db 265 IleLeuSerLeuValAsnAsnAspAspArgAspSerTyArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGAAATCCCATCCGGTTTCATTTCTACATCTTGAACCGCCATGACAAACAGAACCTC 898
||| : : : : : : : : : : :
Db 285 LeuArgValProSerGlyThrThrTyTyTyValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGCTAAATCTCCATGCCCGTTAACACACCGCGCCAGTTTGAGGATTTCTCCCG 958
||| : : : : : : : : : : :
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY 959 GCGAGACCGGAGACCAATCATCTACTTGCAGGCGTTCAGCAGGAATACGTTGAGGCG 1018
: : : : : : : : : : :
Db 325 SerSerThrGluAlaGlnSerTyLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCCGGAATTCATGATACGAGGAGGTCTGTGAAGAGAATCGCAGAGGT 1078
: : : : : : : : : : :
Db 345 SerTyAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GAGCAAGAGAGAGAGGCGAGCGCATGGAGTACTCGGAGTAGTGAGAACCAATGAAGA 1138
||| : : : : : : : : : : :
Db 364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAGAGGAGCAGCTTGAAGAACTTACTAAGCACCGCTAAATCCGTC 1198
||| : : : : : : : : : : :
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394

QY 1199 TCAAAAGAAAGGCTCCGAAAGAGAGGAGATATCACCAACCCCAATCAACTTGAGAGAAGC 1258
||| : : : : :
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTCTAAACAACCTTGGGAAGTTATTGAGTCAAGCCAGACAAGAGAAC 1318
: : : : : : : : : : :
Db 412 AspProIleTySerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGGACCTCGACATGATGCTCACTGTGTAGAGATCAAAAGAGGAGCTTTG 1378
||| : : : : : : : : : : :
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCACACTTCAACTCAAGCCCATGTGTTATCGTCGTCAACAAAGGAAGTGA 1438
: : : : : : : : : : :
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCGTGGCTGTAAGAAAGACCAACAGAGGGGACGGGGGAAGAG 1498
||| : : : : : : : : : : :
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAAGAGGAGGAAAGTAACAGAGAGGTGCGTAGGTACACAGCAGG 1558
||| : : : : : : : : : : :
Db 490 Pro-----LeuGluValArgLysTyArgAlaGlu 499
QY 1559 TTGAAGGAAGCGATGTGTTTCATATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
||| : : : : : : : : : : :
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyProValValValAsnAlaThr 519
QY 1619 TCCGAACCTCCATCTCTGCTTGGCTTCAACGCTGAAACCAACACAGATCTTCTCTT 1678
||| : : : : : : : : : : :
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAGGACAATGTGATAGACCAAGAGTAGAAGCAAGCAAGGATTTAGCATTC 1738
||| : : : : : : : : : : :
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTGGGTCCGGTGAAACAAGTTTGAGAAAGCTCATCAAAAAACCAAGAAATCTCACTTTGTG 1798
||| : : : : : : : : : : :
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyPheVal 579
QY 1799 AGTGTCTCTCTCAATCTCAATCTCAATCTCGTGTCTCTCGAGAAAGAGTCTCTCTGAG 1858
||| : : : : : : : : : : :
Db 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
: : : : : : : : : : :
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
||| : : : : :
Db 603 AlaPhe 604
RESULT 2
FMSYBA
beta-conglycinin alpha chain precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text_change 09-Jul-2004
C:Accession: S14681; S74124; S06714
R:Sebastiani, F.L.; Farrell, L.B.; Schuler, M.A.; Beachy, R.N.
Plant Mol. Biol. 15, 197-201, 1990
A:Title: Complete sequence of a cDNA of alpha subunit of soybean beta-conglycinin.
A:Reference number: S14681; MUID:91355860; PMID:2103438
A:Accession: S14681
A:Molecule type: mRNA
A:Residues: 1-605 <SEB>
A:Cross-references: UNIPROT:P13916; EMBL:X17698; NID:g18535; PIDN:CAA35691.1; PID:g18536
R:Shutov, A.D.; Kakhovskaya, I.A.; Bastrygina, A.S.; Bulmaga, V.P.; Horstmann, C.; Muent
Eur. J. Biochem. 241, 221-228, 1996
A:Title: Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage gl
A:Reference number: S74123; MUID:97054613; PMID:8898910
A:Accession: S74124
A:Molecule type: protein
A:Residues: 189-196, 'H', 198, 'N', 200, 'X', 202-203; 397-408, 'X', 410, 'X', 412-417, 'X', 419-420, 'X'

A; Experimental source: seed

A; Experimental source: seed
C; Superfamily: glycinin
F; 1-22/Domain: signal sequence #status predicted <SIG>
F; 23-62/Domain: propeptide #status predicted <PRO>
F; 63-605/Product: beta-conglyonin alpha chain #status
F; 261, 517/Binding site: carbohydrate (Asn) (covalent)

Alignment Scores:

Pred. No.:	9.43e-82	Length:	605
Score:	1281.00	Matches:	285
Percent Similarity:	59.21%	Conservative:	107
Best Local Similarity:	43.05%	Mismatches:	174
Query Match:	35.96%	Indels:	96
DB:	1	Gaps:	18

US-10-728-323-1 (1-2032) x FWSYBA (1-605)

QY	50	ATGAGAGGAGGTTTCTCCACTGATGCTGTGTCTAGGATCTCTGTCTGGCTTCAGTT	109
DB			
DB	2	MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal	19
QY	110	TCGCAACGCATGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTGCGCCAG	169
DB		:::	
DB	20	SerValSerPheGlyLeuAla-----TyrTrpGluLys--GluAsnProLysHisAsn	36
QY	170	AGGTGCCTCCAGAGTTGTCAACAGCAACCGGATGACTTTGAAGCAAAAGGATCGAGTCT	229
DB			
DB	37	LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56
QY	230	CGCTGCACCAAGCTCGATGATGAT-----CTCGTTGT	262
DB			
DB	57	ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluLeuPheArg--	75
QY	263	GTCTATGATCTCTCGAGGACACACTGGCACCAACCAACGTTCCCTCCAGGGGACGG	322
DB			
DB	76	-----ProArgProArg-----ProGlnHisProGlu	84
QY	323	ACACGTGGCCCAACCCGGAGACTACGATGATGAC-----	358
DB			
DB	85	ArgGluProGlnGlnProGlyGluLysGluGluAspGluGlnProArgProIle	104
QY	359	-----CGCGTCAACCCCGAGAGAG-----GAAGGAGCGCA	391
DB			
DB	105	ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGluGln	124
QY	392	TGG-----GCACGAGCTGCACCGAGGAGCGCTGAAGAGAGAAGAA	430
DB			
DB	125	TrpProArgLysGluLysArgGlyGlyLysGlySerGluGluGluAspGluAspGlu	144
QY	431	GACTGGAGACAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCTATCAGCACGACCG	484
DB			
DB	145	AspGluGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu	164
QY	485	AAATAAGGCCCGCAAGGAAGAGAGGACCAACAGAGTGGGGACACACGATGCCATGTC	544
DB			
DB	165	ArgAsnGluGluLysAspGluAspGluGlnArgGlnArgGluSerGluAsp	184
QY	545	AGGGAA-----GAACAATCTCGGAACAACCTTTCTACTTCCGTCAGGCGGTTTAGC	598
DB			
DB	185	SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu	204
QY	599	ACCCGCTACGGGAACCAAAACGGTAGGATCCGGCTCTCTGCAGAGGTTTGACCAAGGTCA	658
DB			
DB	205	ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer	224
QY	659	AGCGAGTTTTCAGAATCTCCAGAAATCACCGTATTCTGCAGATCTCGAGGCAAACTTAAC	718
DB			
DB	225	ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr	244
QY	719	CTTGTCTTCCCAAGCAGCGCTGATGATTAACATCTCTGTTATCTCAGCAAGGGCAAGCC	778
DB			
DB	245	LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla	264

Qy 1505 GAAGACGAAGAAGAGGAGGAAGTAACAAGAGAGGTGGTGGTAGGTATACAGCAGCGAGGTGGAAG 1564
||||| :||| ||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 450 GluAspArgLysGluAArg-----AsnAsnGluValGlnArgTyrGluAlaAArgLeuSer 467
||||| :||| ||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1565 GAAGCCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTACGCATCAAGCTTCTCTCCGAA 1624
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 468 ProGlyAspValvallelleProallaglyhisProvalalalleserAlaserSerAsn 487
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1625 CTCATCTGCTTGCTTCGCTATCAACGCTCAAACCAACCACACAGAAATCTTCTTCGAGGT 1684
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 488 LeuAsnLeuLeuGlyPheGlyIleAsnAlalyAsnAsnGlnArgAsnPheLeuSerGly 507
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1685 GATAAGACACATGTGATAGACCATAGAGAGAACGACGAGGATTTAGCATTTCCCTGGG 1744
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 508 SerAspAspAsnVallelleserGlnleGluAsnProVallysGluLeuThrPheProgly 527
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1745 TCGGGTCAACAAGTTGAGAGCTCATCAAAAACAGAGGAGGAATCTCACTTTGTCAGTGT 1804
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 528 SerSerGlnGluAlaAsnAArgleulleGlyAsnGlnlysGlnserHisPheAlaserAla 547
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1805 CGTCCTCAATCTCAATCTCAATCTCCGCTCTCTGAGAAAGAGTCTCTCGAGAAAGAG 1864
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 548 GluPro-----Glu 550
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 1865 GATCAAGAGAGGAAACCAAGGAGGAGGTCCTCTTCAATTTTGAAGGCTTTT 1924
:||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 551 GlnlysGluGluSerGlnArgLysArgSerProLeuSerValLeuAspSerPhe 570
:||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||

RESULT 4

B24810
beta-conglycinin alpha chain precursor - soybean
C;Species: Glycine max [soybean]
C;Date: 07-Mar-1988 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C;Accession: B24810; S16337
R;Doyle, J.J.; Schuler, M.A.; Godette, W.D.; Zenger, V.; Beachy, R.N.; Slightom, J.L.
J. Biol. Chem. 263, 9228-9238, 1988
A;Title: The glycosylated seed storage proteins of Glycine max and Phaseolus vulgaris.
A;Reference number: A24810; PMID:86250867; PMID:3013879
A;Accession: B24810
A;Molecule type: DNA
A;Residues: 1-639 <DOY>
A;Cross-references: UNIPROT:P11827; GB:M13759; NID:g169928; PIDN:AAB01374.1; PID:g169929
A;Note: the authors translated the codon GGT for residue 352 as Glu
R;Schuler, M.A.; Schmitt, E.S.; Beachy, R.N.
Nucleic Acids Res. 10, 8225-8244, 1982
A;Title: Closely related families of genes code for the alpha and alpha' subunits of the
A;Reference number: S16337; PMID:8314288; PMID:6298713
A;Accession: S16337
A;Molecule type: DNA
A;Residues: 361-639 <SCH>
C;Genetics:
A;Introns: 278/1; 355/3; 382/3; 481/3; 575/1
C;Superfamily: glycinin
C;Keywords: seed; storage protein

Alignment Scores:

Pred. No.:	1-28-79	Length:	639
Score:	1251.00	Matches:	283
Percent Similarity:	58.75%	Conservative:	120
Best Local Similarity:	41.25%	Mismatched:	173
Query Match:	35.12%	Indels:	110
DB:	2	Gaps:	21

US-10-728-323-1 (1-2032) x B24810 (1-639)

Qy 50 ATGACAGGAGGGTTTCTCCACTGATCTGTGTAGCGATCTCTGCTGGTTCAGATT 109
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuGlyValPheLeuAlaSerVal 19
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Qy 110 TCTGCACGATGCGCAGTCAATCACCTTACCAGNAGAAACAGAGACCCCTCGGCCAG 169
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36
||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :||| ||||| :|||

```
Db      397  |||  ::|||  |||||  |||  407
      GluGlycInGlnGlnGluGluArgLeuGln-----
Qy      1127 AACAAATCAGGAGTGATGCTCAAGTGTCAAGGACGCTTCAAGAACTTACTAGCAC 1186
Db      408 -----GluSerValIleValGluIleSerLysLysGlnIleArgGluLeuSerLysHis 425
Qy      1187 GCTAAATCCCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCAATCAAC 1246
Db      426 AlaLysSerSerArgLysThrIleSerSerGlu-----AspLysProPheAsn 442
Qy      1247 TTGAGAGAAGCGAGCCGATCTTTCTAAACAATTTGGGAAGTATTATTCAGGTGAAGCCA 1306
Db      443 LeuGlySerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIle----- 460
Qy      1307 GACAAAGNAGACCCAGCTTCAGGACCTCGACATGATCTCACCTGTGTAGATGACAA 1366
Db      461 ThrGlnArgAsnProGlnLeuArgAspLeuAspValPheLeuSerValValAspMetAsn 480
Qy      1367 GAAGGAGCTTTGATGCTCCACACTTCAACTCAAAAGCCATGTTATCTCGTGTGCTCAAC 1426
Db      481 GluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValIleAsn 500
Qy      1427 AAGGAACTCGAAACCTTGAACTCGTGTGGTGTAAAGAAAGACAAACAGAGGGGACGG 1486
Db      501 GluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGlnGln 519
Qy      1487 CGGAGNAGAGGAGGACGACGACGAGAGAGAGGAGGAGTAACAGAGAGTGGTAGG 1546
Db      520 GlnGluGluGlnPro-----LeuGluValArgLys 529
Qy      1547 TACACAGCGAGGTTGAAGAGCGGATGTTTCATCATGCCAGCCTCATCCAGTAGCC 1606
Db      530 TyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValMet 549
Qy      1607 ATCAACGCTTCCTCGAACTCCATCTGCTTGGCTTCGGTATCAAGCTCAAAACCAACCA 1666
Db      550 ValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsnGln 569
Qy      1667 AGAATCTCTCTGCAGGTATAGGACAAATGTTGAGAAGCTCATCAAAACCAAGAGAA 1786
Db      590 GluLeuAlaPheProArgSerAlaLysAspIleGluAsnLeuIleLysSerGlnSerGlu 609
Qy      1787 TCTCACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGCTCCTCGAGAAA 1846
Db      610 SerTyrPheValAspAlaGlnProGln----- 618
Qy      1847 GAGTCTCCTGAGAAAGGATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1906
Db      619 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeuSer 632
Qy      1907 TCAATTTTGAAGGCTTTT 1924
Db      633 SerIleLeuArgAlaPhe 638

RESULT 5
S08505
vicilin - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S08505
R:Higgins, T.J.V.; Newbigin, E.J.; Spencer, D.; Llewellyn, D.J.; Craig, S.
A:Title: The sequence of a pea vicilin gene and its expression in transgenic tobacco pla
Plant Mol. Biol. 11, 683-695, 1988
A:Reference number: S08505
A:Accession: S08505
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <HIG>
```

A;Cross-references: UNIPROT:P13918; EMBL:X14076; NID:g20917; PIDN:CAA32239.1; PID:g313671
C:Genetics:
A:Introns: 102/1; 160/3; 187/3; 295/3; 400/1
C:Superfamily: glycinin

Alignment Scores:
Pred. No.: 1.33e-74 Length: 459
Score: 1178.50 Matches: 223
Percent Similarity: 71.33% Conservative: 93
Best Local Similarity: 50.34% Mismatches: 114
Query Match: 33.09% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S08505 (1-459)

```
Qy      542 GTGAGGGAGAGAACATCTCGGNAACACCGCTTCTACTTCCCGTCAAGCGGTTTAGCACC 601
Db      26 ValSerSerArgSerAspProGlnAsnProPheIlePheLysSerAsnLysPheGlnThr 45
Qy      602 CGCTACGGGAACCAAAAACGCTAGGATCGGGTCTCGACAGAGGTTTGACCAAGGTCAAGG 661
Db      46 LeuPheGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnArgSerLys 65
Qy      662 CAGTTTCAGATCTCCAGAATCACGATATGTGCAGATCGAGGCCAAACCTTAACACTCTT 721
Db      66 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIle 85
Qy      722 GTTCTTCCAGCAGCCTGATGCTGATACATACCTCTGTATTCACGACGAGGCAAGCCACC 781
Db      86 PheLeuProGlnHisThrAspAlaAspTyrIleLeuValValLeuSerGlyLysAlaIle 105
Qy      782 GTGACCGTAGCAATGCGCAATAACAGAAAGAGCTTAACTCTTGACGAGGCCATGCATC 841
Db      106 LeuThrValLeuLysProAspAspArgAsnSerPheAsnLeuGluArgGlyAspThrIle 125
Qy      842 AGAATPCCATCCGTTTCATTTCTTACATCTTGAAACCGCCATGACAAACAGAACCTCAGA 901
Db      126 LysLeuProAlaGlyThrIleAlaTyrLeuValAsnArgAspAsnGluGluLeuArg 145
Qy      902 GTAGCTAAATCTCCATCCGCTTAAACACACCGCCGCTTTGAGGATTTCTCCCGGCG 961
Db      146 ValLeuAspLeuAlaIleProValAsnArgProGlyGlnLeuGlnSerPheLeuLeuSer 165
Qy      962 AGCAGCCGAGACCAATCATCTTTCAGGCGCTTCAGCAGGAAATACGTTGAGGCGCCGC 1021
Db      166 GlyAsnGlnAsnGlnAsnTyrLeuSerGlyPheSerLysAsnIleLeuGluAlaSer 185
Qy      1022 TTCATTCGGGAATTCOAATGAGATACGAGGCGTGTGTAGAGAGAATGCGAGGAGTGAG 1081
Db      186 PheAsnThrAspTyrGluGluIleGluLysValLeuLeuGluGluHisGluLysGluThr 205
Qy      1082 CAAGAGGAGAGGCGGATGAGTACTCGGAGTAGTGAGAACAACTCAAGGAGTG 1141
Db      206 GlnHisArgGserLeuLys-----AspLysArgGlnGlnSerGlnGluAsnVal 223
Qy      1142 ATAGTCAAAGTGTCAAAGGAGCAGCTTCAAGAACTTACTTAAGCAGCTATAATCCCTCTCA 1201
Db      224 IleValLysLeuSerArgGlyGlnIleGluLeuSerLysAsnAlaLysSerThrSer 243
Qy      1202 AAGAAAGGCTCCGAAGAAGAGGAGGATATCAACCAACCAATCAACTTGAGAGAGGCGAG 1261
Db      244 LysLysSerValSerSerGluSerGlu-----ProPheAsnLeuArgSerArgGly 260
Qy      1262 CCCGATCTTCTTACAACTTTGGGAAGTTATTTCAGGTGAAGCCAGACAGAGAACCC 1321
Db      261 ProIleTyrSerAsnGluPheGlyLysPhePheGluIleThrProGlu---LysAsnPro 279
Qy      1322 CAGCTTCAGGACCTCGACATCATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATG 1381
Db      280 GlnLeuGlnAspLeuAspIlePheValAsnSerValGluIleLysGluGlySerLeuLeu 299
Qy      1382 CTCACACACTTCAACTCAAGGCCCATGTTTATCGTCGTCGTCACAAAGAGAACTCGAAAC 1441
```

[illegible]

Qy	746	GATAAATCATCTTGTATTATCCAGCAAGGGCAAGCCACCGTGACCGCTAGCACAATGTCGAATAAC	805
Db		::: ::: ::: ::: ::: :::	
	68	AspTyrIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeuLysProAspAsp	87
Qy	806	AGAAAGAGCTTTTAATCTTGACGAGGGGCATGCACTCAGAAATCCCATCCGGTTCATTTCC	865
Db		::: ::: ::: ::: ::: :::	
	88	ArgAsnSerPheAsnLeuGluArgGlyAspThrIleLysLeuProAlaGlyThrIleAla	107
Qy	866	TACATCTGAACCGCCATGACCAACAGAACCTCAGAGTAGCTAAAAATCTCCATGCCGCTT	925
Db		::: ::: ::: ::: ::: :::	
	108	TyrIeuValAsnArgAspAsnGluLuleuArgValLeuAspLeuAlaIleProVal	127
Qy	926	AACACACCCCGCCAGTTTGAGGATTTCTTCCC CGCAGCAGCCGAGACC AATCATCTCAT	985
Db		::: ::: ::: ::: ::: :::	
	128	AsnArgProGlyGlnLeuGlnSerPheLeuLeuSerGlyAsnGlnAsnGlnAsnTyr	147
Qy	986	TTCCAGGGCTTCAGCAGGATACTGTTGGAGCGCCTTCAAATCCCGGNATTCANTAGATA	1045
Db		::: ::: ::: ::: ::: :::	
	148	LeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAspTyrGluGluIle	167
Qy	1046	CGGAGGTCGTGTTAGAAGAAATG CAGGAGGTGTGACAGAGAGAGAGCGCAGAGCGGA	1105
Db		::: ::: ::: ::: ::: :::	
	168	GluLysValLeuLeuGluGluHisGlyLysGluThrGlnHisArgArgSerLeuLys---	186
Qy	1106	TGAGTACTCGGAGTAGTGAGAA CAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGCAC	1165
Db		::: ::: ::: ::: ::: :::	
	187	--AspLysArgGlnGlnSerGlnGluGluAsnValIleValLysLeuSerArgGlyGln	205
Qy	1166	GTTGAAGAACTTACTAAGCAGCGCTAAATCCGTC TCAAAGAAAGGCTCCGAAGAAGAGGGA	1225
Db		::: ::: ::: ::: ::: :::	
	206	IleGluLeuSerLysAsnAlaLysSerThrSerLysLysSerValSerSerGluSer	225
Qy	1226	GATATCACCAACCCAATCAACTTTGAGAGAGGCGAGCCCGATCTTTCTAACAACTTTGGG	1285
Db		::: ::: ::: ::: ::: :::	
	226	Glu-----ProPheAsnLeuArgSerArgGlyProIleTyrSerAsnGluPheGly	242
Qy	1286	AAGTTATTTGAGTGAAGCCAGACAGAGAAGACCCCGAGCTTCAGGACCTGGACATGATG	1345
Db		::: ::: ::: ::: ::: :::	
	243	LysPhePheGluIleThrProGlu---LysAsnProGlnLeuGlnAspLeuAspIlePhe	261
Qy	1346	CTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGTAGTCTCCACACTTCAACTCAAAGGCC	1405
Db		::: ::: ::: ::: ::: :::	
	262	ValAsnSerValGluLeuLysGluLysLeuLeuLeuLeuProHisTyrAsnSerArgAla	281
Qy	1406	ATGTTTATCGTCTGCTCAACAAAGGAATGGNAACCTTGAACCTCTGTGCTGTGAAGAAA	1465
Db		::: ::: ::: ::: ::: :::	
	282	IleValIleValThrValAsnGluLysGlyAspPheGluLeuValGlyGlnArgAsn	301
Qy	1466	GAGCAACACAGAGGGGACGGCGGGGAAGAGGAGGACGAGACCAAGAGGAGGGGA	1525
Db		::: ::: ::: ::: ::: :::	
	302	GluAsnGlnGlnGluGlnArgLysGluAspAspGluGluGlnGlnGlyGluGluGlu	321
Qy	1526	AGTAACAGAGAGTGGCTAGGTATACACGCGAGGTTTGAAGGAAGGCCATGTGTTTCATCATG	1585
Db		::: ::: ::: ::: ::: :::	
	322	IleAsnLysGlnValGlnAsnTyrLysAlaLysLeuSerSerGlyAspValPheValIle	341
Qy	1586	CBAGAGCTCATCCAGTAGCCATCAACGGTTCTCTCGAACTCCATCTGCTTGGCTTCGGT	1645
Db		::: ::: ::: ::: ::: :::	
	342	ProAlaGlyHisProValAlaLeuLysAlaIaSerSerAsnLeuAspLeuGlyPheGly	361
Qy	1646	ATCAACGCTGAAACCAACACAGAACTCTTCTTGCAGGTGATAAGACAAATGTGATAGAC	1705
Db		::: ::: ::: ::: ::: :::	
	362	IleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlyAspGluAspAsnValIleSer	381
Qy	1706	CAGATAGAGACCAACGCAAGGATTTAGCATTTCCCTGGGTCGGGTGAACCAAGTTGAGAAG	1765
Db		::: ::: ::: ::: ::: :::	
	382	GlnValGlnArgProValLysGluLeuAlaPheProGlySerAlaGlnGluValAspArg	401
Qy	1766	CTCATCAAAAACAGAAAGAAATCTCATTTGTGAGTGTCTGCTCAATCTCAATCTCAA	1825
Db		::: ::: ::: ::: ::: :::	
	402	IleLeuGluAsnGlnLysGlnSerHisPheAlaAspAlaGlnProGlnGlnArg-----	419

QY 1826 TCTCCGTCGTCCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAA 1870
 Db 420 -----GluArgGlySerArgGluThrArgAspArg 429
 RESULT 7
 FWSYCB
 beta-conglycinin beta chain - soybean
 C:Species: Glycine max (soybean)
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
 C:Accession: JQ0969
 R:Harada, J.J.; Barker, S.J.; Goldberg, R.B.
 Plant Cell 1, 415-425, 1989
 A:Title: Soybean beta-conglycinin genes are clustered in several DNA regions and are regulated by phytochrome
 A:Reference number: JQ0969; MUID:93005638; PMID:2562562
 A:Accession: JQ0969
 A:Molecule type: DNA
 A:Residues: 1-439 <HAR>
 A:Cross-references: UNIPROT:P25974; GB:S44893; NID:G256426; PIDN:AB23463.1; PID:G256427
 C:Comment: This protein accumulates during seed development and is hydrolyzed after germination
 C:Genetics:
 A:Introns: 10/1; 159/3; 186/3; 281/3; 375/1
 C:Superfamily: glycinin
 C:Keywords: seed; storage protein
 Alignment Scores:
 Pred. No.: 4.95e-73 Length: 439
 Score: 1156.00 Matches: 237
 Percent Similarity: 69.98% Conservative: 87
 Best Local Similarity: 51.19% Mismatches: 85
 Query Match: 32.45% Indels: 54
 DB: 1 Gaps: 10
 US-10-728-323-1 (1-2032) x FWSYCB (1-439)
 QY 542 GTGAGGAGAAACATCTCGGAACAACCCCTTTCTACTTC---CCGTCAAGCGGTTTACG 598
 Db 26 ValArgGluAsp-----GluAsnAsnProPheTy-PheArgSerSerAsnSerPheGln 43
 QY 599 ACCCGCTACGGGAACCAAAACGGTAGATCCGGTCTCGAGAGGTTTACCAAAAGGTCA 658
 Db 44 ThrLeuPheGluAsnGlnAsnValArgIleArgLeuLeuGlnArgPheAsnLysArgSer 63
 QY 659 AGGCAGGTTTCAGAAATCTCCAGAATCAACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
 Db 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 83
 QY 719 CTGTGTTCTCCCAAGACGCTGATGTGATAACATCTCTGTTATCCAGCAAGGGCAAGCC 778
 Db 84 IleLeuLeuProHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 103
 QY 779 ACCGTGACGTGACAAATGCAATACAGAAACAGCTTTAATCTTCGAGGGCCCATGCA 838
 Db 104 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 123
 QY 839 CTGAGATCCCATCCCGGTTTCATTTCCTACATCTTCAACCGCCATCAACACCAAGACCTC 898
 Db 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
 QY 899 AGAGTAGCTAAATCTCCATGCGCGTTTAAACACACCCGCGCAGTTTGAGGATTTCTTCCCG 958
 Db 144 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 163
 QY 959 GCGAGGACGCGAGACCAATCATCTCTACTGCGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
 Db 164 SerSerGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnLeuLeuGluThr 183
 QY 1019 GCCTTCAATCCGGAATTCATAGATACGAGGGGTGCTGTAGAGAGAATGCAGAGGT 1078
 Db 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 198
 QY 1079 GACCAAGAGGAGAGGGCGAGCGCATGGAGTACTCTCGGAGTAGTGAGAACATGAAGA 1138
 Db 1138 -----Gly 198

Alignment Scores:

Pred. No.: 1,8e-72 Length: 438
Score: 1148.00 Matches: 222
Percent Similarity: 73.58% Conservative: 90
Best Local Similarity: 52.36% Mismatches: 100
Query Match: 32.23% Indels: 12
DB: 2 Gaps: 5

US-10-728-323-1 (1-2032) x S35757 (1-438)

```
QY 542 GTGAGGAGAAACATCTCGAACAACCCCTTCTACTTCCGTCAGGCGGTTTAGCACC 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 22 ValSerSerArgSerAspGlnGluuAsnProPheIlePheLysSerAsnArgPheGlnThr 41

QY 602 CGTACGGGACCAAAACCGTAGATCCCGGTCTCGAGAGTTTGACCAAGGTCAAGG 661
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 42 LeuTyrgluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 61

QY 662 CAGTTTCAGATCTCCAGATCACCGTATGTGTCAGATCGAGCCCAACCTAACACTCTT 721
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 IlePheGluAsnLeuGlnAsnTyrgluLeuLeuGlyTyrgluSerLysProHisThrLeu 81

QY 722 GTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAAGCGCAAGCCACC 761
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 PheLeuProGlnTyrgluThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaThr 101

QY 782 GTGACCGTAGCAATCGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCCATGCAC 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 121

QY 842 AGAATCCCATCCGGTTTCTTCTCATCTTCACTTGAACCGCATCAGCAACCAACCTCAG 901
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 LysLeuProAlaGlyThrIleAlaTyrgluLeuAlaAsnArgAspAsnGluAspLeuArg 141

QY 902 GTAGCTAAATCTCCATGCCCGCTTAACACACCCGCGCAGTTTCAGGATTTCTCCCGCG 961
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 161

QY 962 AGCAGCGGAGACCAATCATCTTCTACTTTCAGGGCTTCAGCAGGAAATACGTTGAGCG 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 GlyThrGlnAsnGlnProSerLeuLeuSerGlyPheSerLysAsnIleLeuGluAlaIle 181

QY 1022 TTCAATCGGATTCATGATGATACGAGGCTCTGTAGACAGAAATCGCAGAGGTGAG 1081
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 PheAsnThrAsnTyrgluGluIleGluValLeuLeuGluGln-----Glu 198

QY 1082 CAAGAGGAGAGGCGCAGAGCGCATCGAGTACTCGAGTAGTGAG---AACATGAAGGA 1138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGluLeuAsnGluGluAsn 218

QY 1139 GTGATAGTCAAGTGTCAAAGGAGCAGCTTTGAAGAACTTACTAAGCAGCGCTAAATCCG 1198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 ValIleValLysValSerArgGluGlnIleGluLeuSerLysAsnAlaLysSerSer 238

QY 1199 TCAAGAAAGCGTCCGAAGAGGAGGATATACCAACCAATCACTTGTAGAGAGGC 1258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 255

QY 1259 GAGCCCGATCTTTCAACAACTTTGGAGAGTATTTGAGTGAGCGCAGCAGAGAGAAC 1318
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 256 AsnProIleTyrgluSerAsnLysPheGlyLysPhePheGluIleThrProGlu---LysAsn 274

QY 1319 CCCAGCTTCAGGACCTGAGCATGATGCTCACCTGTGTAGAGATCAAGAGAGGAGCTTG 1378
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 275 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysGluGlySerLeu 294

QY 1379 ATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAAGAACTCGGA 1438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 LeuLeuProAsnTyrgluSerArgAlaIleValIleValThrValThrGluGlyLysGly 314

QY 1439 AACCTTGAATCTGTGCTGTGAAGAAAGACCAACAGAGGCGCGGGAAGAG 1498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 315 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 330
QY 1499 GAGGACGAAGACGAAGAGAGGAGGAGGAGTAACAGAGAGGTGCTAGGTACACAGCAGG 1558
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 AspLysGluGluGluGlnGluGluThrSerLysGlnValGlnLeuTyrgluAlaLys 350
QY 1559 TTGAAGGAGGCGATGTTTCATCGCCAGAGCTCATCCAGTACCGCATCAACGCTTCC 1618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 370
QY 1619 TCCGAACCTCATCTGCTCGCTTCCGATCAACGCTGAAACCAACCAACCAAGATCTTCC 1678
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 371 SerAspLeuAsnLeuIleGlyPheGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeu 390
QY 1679 GAGGTGATTAAGACAATGTGATAGACCATAGAGAACAAAGCAGGAGGATTTAGCATTC 1738
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 410
QY 1739 CTTGGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAACAGAGAAATCTCACTTTGTG 1798
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 ProGlySerSerHisGluValAspArgLeuLeuLysAsnGlnLysGlnSerTyrgluAla 430
QY 1799 AGTGTCTCGTCT 1810
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 431 AsnAlaGlnPro 434
```

RESULT 9

A27288

vicilin precursor - fava bean

C:Species: Vicia faba (fava bean)

C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004

C:Accession: A27288

R:Basuener, R.; Van Hai, N.; Jung, R.; Saalbach, G.; Muentz, K.

Nucleic Acids Res. 15, 9609, 1987

A:Title: The primary structure of the predominating vicilin storage protein subunit from

A:Reference number: A27288; MUID:88067789; PMID:3684610

A:Accession: A27288

A:Molecule type: mRNA

A:Residues: 1-463 <BAS>

A:Cross-references: UNIPROT:P08438; GB:Y00462; NID:g22052; PIDN:CAA68525.1; PID:g22053

C:Superfamily: glycinin

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-463/Product: vicilin #status predicted <VCN>

Alignment Scores:

Pred. No.: 9e-70 Length: 463
Score: 1109.50 Matches: 218
Percent Similarity: 68.93% Conservative: 97
Best Local Similarity: 47.70% Mismatches: 119
Query Match: 31.15% Indels: 23
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x A27288 (1-463)

```
QY 554 ACATCTCGAACAACCCCTTCTACTTCCGTCAGGCGGTTTAGCACCCTAGCGGAAC 613
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 30 SerAspGlnAspAsnProPheValPheGluSerAsnArgPheGlnThrLeuPheGluAsn 49
QY 614 CAAAACGGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCAAGGCGAGTTTTCAGAAT 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50 GluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnHisSerLysLeuLeuGluAsn 69
QY 674 CTCAGAATCACCGTATTGTGCAGATCGAGGCCAACCTTAACACTCTTCTTCTCCCAAG 733
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 LeuGlnAsnTyrgluLeuLeuGluTyrgluSerLysProHisThrIlePheLeuProGln 89
QY 734 CACGTGTATGATTAACATCTTGTATCCAGCAGGCGCAGCCACCGTACCGGTAGCA 793
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 GlnThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaIleLeuThrValLeu 109
QY 794 AATGCAATTAACAGAAAGAGCTTTAATCTTTCAGAGGCGCATGCATCAGAAATCCCATCC 853
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 LeuProAsnAspArgAsnSerPheSerLeuGluArgGlyAspThrIleLysLeuProAla 129
```

```
QY 854 GGTTCATTCTTACATCTTGAACCGCCATGACCAACGAGAACCTCAGAGTAGCTAAAATC 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyThrIleGlyTyrLeuValAsnArgAspGluAspLeuValLeuAspLeu 149

QY 914 TCATGCCCGTTAAACACACCGGCCMTGTTGAGGATTTCTTCGCGGAGACGCGAGAC 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ValIleProValAsnArgProGlyGluProGlnSerPheLeuLeuSerGlyAsnGlnAsn 169

QY 974 CAATCATCTACTTGCAGGCGTTTCAGCAGGATACGTTGAGGCGCCCTTCATCGGAA 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlnProSerIleLeuSerGlyPheSerIysAsnIleLeuGluAlaSerPheAsnThrAsp 189

QY 1034 TTCAATGATACGAGGAGTGTCTTTAGAAAGAGATTCAGGAGGTGCACAGCAGAGAGA 1093
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db TyrLysGluIleGluLysValLeuLeuGluGluHisGlyLysGluLysTyrHisArgArg 209

QY 1094 GGGCAGAGCGGATGGAGTACTCGGAGTAGTAGAACATGAAGAGTAGTAGTCAAAGTG 1153
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyLeuLysAspArgArgGlnArgGlyGlnGluGluAsn-----ValIleValIlystle 227

QY 1154 TCAGAAGGACACGTTGAAGCAACTTACTAAGCACGCTAAATCGTCTCAAAGAAAGCTCC 1213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db SerArgLysGlnIleGluLeuAsnLysAsnAlaLysSerSerSerLysLysSerThr 247

QY 1214 GAAGAAGAGGAGATATATCAACCAACCAATCAACTTGAGAGAGCGGAGCCGATCTTTCT 1273
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db SerSerGluSerGlu-----ProPheAsnLeuLysSerArgGluProIleTyrSer 264

QY 1274 ACAACTTTGGAAAGTTATTGAGGTGAAGCCAGACAGAAACCCCGAGTTTCAGGAC 1333
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnLysPheGlyLysPhePheGluIleThrPro---LysArgAsnProGlnLeuGlnAsp 283

QY 1334 CTGGACATCATCTCACCTGTCTAGAGATCAAGAAGGAGCTTGTATGCTCCCACTTC 1393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LeuAsnIlePheValAsnTyrValGluIleAsnGluGlySerLeuLeuLeuProHisTyr 303

QY 1394 AACTCAAAGGCCATGGTTATCGTCTGCTCAACAAAGGAACCTGGAAACCTTGAACCTCGTG 1453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323

QY 1454 GCTGTAAGAAAGACMAACACAGAGGGGACGGCGGGGAAGAGAGAGACGACGACGAA 1513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyGlnArgAsnGluAsnGlnGlnGlyLeuArgGluGluTyrAspGluGlyLysGln 343

QY 1514 GAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGTACACAGCGAGGTGTGAAGGAAGCGCAT 1573
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyGluGluIleArgLysGlnValGlnAsnTyrLysAlaLysLeuSerProGlyAsp 363

QY 1574 GTGTTTCATCCGACGAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTG 1633
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ValLeuValIleProAlaGlyTyrProValAlaIleLysAlaSerSerAsnLeuAsnLeu 383

QY 1634 CTTCGCTTCGGTATCAACGCTGAAACACACAGAAATCTTCCTTGACGATGAAGAC 1693
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ValGlyPheGlyIleAsnAlaGluAsnGlnArgTyrPheLeuAlaGlyGluAsp 403

QY 1694 AATGTGATAGACCATAGAGAAGCAAGCAAGGATTTAGCATTCCTGGGTCGGGTCAA 1753
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db AsnValIleSerGlnIleHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423

QY 1754 CAAGTTGAGAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAA 1813
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProGln 443

QY 1814 TCTCAATCTCAATCTCCGTCGTCTCCTGAGAAGAGTCTCCTGAGAAGAGATCAAGAG 1873
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db -----GluArgGlu 446

QY 1874 GAGGAAACCAAGGAGGAGGTCACCTCTTTCATATTTTGAAGGCTTTT 1924
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ArgGlySerGlnGluIleLysAspHisLeuTyrSerIleLeuGlySerPhe 463
```

RESULT 10

```
S06309
vicilin precursor (clone VfVicl) - tick bean
N:Alternate names: 7S seed storage protein
C:Species: Vicia faba var. minor (tick bean)
C:Date: 31-Mar-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: S06309; S06456
R:Weschke, W.; Baeumlein, H.; Wobus, U.
Nucleic Acids Res. 15, 10065, 1987
A:Title: Nucleotide sequence of a field bean (Vicia faba L. var. minor) vicilin gene.
A:Reference number: S06309; PMID:88096511; PMID:3697075
A:Accession: S06309
A:Molecule type: DNA
A:Residues: 1-463 <WES>
A:Cross-references: UNIPROT:P08438; EMBL:Y00506; NID:G829146; PIDN:CAA68559.1; PID:G2205;
A:Note: the authors translated the codon TCT for residue 296 as Cys
R:Weschke, W.; Baesener, R.; van Hai, N.; Crihal, A.; Baeumlein, H.; Wobus, U.
Biochem. Physiol. 183, 233-242, 1988
A:Title: The structure of a Vicia faba vicilin gene.
A:Reference number: S06456
A:Accession: S06456
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-463 <WES2>
A:Note: 257-Lys and 443-His were also found
C:Genetics:
C:Introns: 102/1; 160/3; 187/3; 295/3; 400/1
C:Superfamily: glycinin
C:Keywords: seed; storage protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-463/Product: vicilin #status predicted <MAT>

Alignment Scores:
Pred. No.: 1,06e-69 Length: 463
Score: 1108.50 Matches: 218
Percent Similarity: 68.93% Conservative: 97
Best Local Similarity: 47.70% Mismatches: 119
Query Match: 31.12% Indels: 23
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S06309 (1-463)

QY 554 ACATCTCGGAACAACCCCTTTCTATCTCCGTCAAGCGGTTTAGCACCCTGACCGGAAC 613
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db SerAspGlnAspAsnProPheValPheGluSerAsnArgPheGlnThrLeuPheGluAsn 49

QY 614 CAARACGCTAGGATCCGGTCTCTCAGAGGTTGACCAAGGTCGAAGGCAGCTTTTCAGAT 673
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GluAsnGlyHisIleArgLeuLeuGlnLysPheAspGlnHisSerLysLeuLeuGluAsn 69

QY 674 CTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTTGTCTTCCCAAG 733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrIlePheLeuProGln 89

QY 734 CAGCTGATGCTGATAACATCTTGTATCCAGCAAGGGCAAGCCACCGTGACCGTAGCA 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlnThrAspAlaAspPheIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeu 109

QY 794 AATGCAATACAGAAAGAGCTTTAATCTTCACGAGGCCCATGCATCAGATCCCATCC 853
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db LeuProAsnAspArgAsnSerPheSerLeuGluArgGlyAspThrIleLysLeuProAla 129

QY 854 GGTTCATTTCTCATCTTGAACCGCCATGACAAACCAAGAACCTCAGAGTAGTAAAAATC 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlyThrIleGlyTyrLeuValAsnArgAspGluGluAspLeuArgValLeuAspLeu 149

QY 914 TCCATGCCCGTTAAACACACCGGCCAGTTTGAAGATTTCTCCCGCGAGCAGCGAGAC 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ValIleProValAsnArgProGlyGluProGlnSerPheLeuLeuSerGlyAsnGlnAsn 169

QY 974 CAATCATCTACTTGCAGGCGTTCAGCAGGATACGTTGAGGCGCCCTTCAATCGGAA 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db GlnProSerIleLeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAsp 189
```

```
QY 1034 TTCAATGAGATACGGAGGTCCTGTAGAGAGAAATGCAGGAGTGTAGCAAGAGGAGAGA 1093
Db 190 TyrlsGluileGluLysValLeuLeuGluGluHisGlyLysGluLysTyrlHisArgArg 209
QY 1094 GGCAGAGCGCATCGAGTACTCGAGTAGTGAGAACAAATGAAGGAGTGATAGTCAAAAGTG 1153
Db 210 GlyLeuLysAspArgGlnArgGlyGlnGluLysn-----ValIleValLysIle 227
QY 1154 TCAAGAGGACGCTTGAAGAACTTACTAGACAGCTAAATCCCTCTCAAGAAAGGCTCC 1213
Db 228 SerArgLysGlnIleGluLeuLeuAsnLysAsnAlaLysSerSerLysLysSerThr 247
QY 1214 GAAGAGAGGAGATATACCAACCCAACTCACTTGAGAGAGCGGCCGATCTTTCT 1273
Db 248 SerSerGluSerGlu-----ProPheAsnLeuArgSerArgGluProIleTyrlSer 264
QY 1274 AACAACTTTGGGAGGTATTGTAGGTGAGCGACAGAGAAACCCAGCTTCCAGGAC 1333
Db 265 AsnLysPheGlyLysPhePheGluIleThrPro---LysArgAsnProGlnLeuGlnAsp 283
QY 1334 CTGCAATGATGTCACTCTGTAGATCAAAAGAGGAGCTTTGATGTCCACACTTC 1393
Db 284 LeuAsnIlePheValAsnTyrlValGluIleAsnGluGlySerLeuLeuLeuProHisTyrl 303
QY 1394 AACTCAAGGCCATGTTATCGTCTGCTCAACAAAGGAAGTCCGAACTTGAACCTCGTG 1453
Db 304 AsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323
QY 1454 GCTGTGAAGAAAGAGCAACAAAGAGGGGCGCGGAGAGAGAGGAGGAGGAGGAGAA 1513
Db 324 GlyGlnArgAsnGluAsnGlnGlnGlyLeuArgGluGluTyrlAspGluGluLysGluGln 343
QY 1514 GAAGAGGAGGAGAAATGAACAGAGAGTGCCTGAGTACACAGCGAGGTGAAGGAGCGAT 1573
Db 344 GlyGluGluGluIleArgLysGlnValGlnAsnTyrlLysAlaLysLeuSerProGlyAsp 363
QY 1574 GTGTTCAATGTCAGCAGCTCATCCAGTAGCCATCAAGCTTCTCCGAACTCCATCTG 1633
Db 364 ValLeuValIleProAlaGlyTyrlProValAlaIleLysAlaSerSerAsnLeuAsnLeu 383
QY 1634 CTGGCTTCGTTCAACGCTGAAGCAACCAACCAAGCAATCTCTTCAGGTGATAGGAC 1693
Db 384 ValGlyPheGlyIleAsnAlaGluAsnAsnGlnArgTyrlPheLeuAlaGlyGluLysAsp 403
QY 1694 AATGTGATAGACCATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGTGCGGTGAA 1753
Db 404 AsnValIleSerGlnIleHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423
QY 1754 CAAGTTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCAA 1813
Db 424 GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProArg 443
QY 1814 TCTCAATCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAG 1873
Db 444 -----GluArgGlu 446
QY 1874 GAGGAAACCAAGAGGAGGAGGCTCCACTCTTTCAATTTGAAGGCTTTT 1924
Db 447 ArgGlySerGlnGluIleLysAspHisLeuTyrlSerIleLeuGlySerPhe 463
```

RESULT 11

```
FWPMVB
vicilin B precursor - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C>Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03344
R:Lycett, G.W.; Delauney, A.J.; Gatehouse, J.A.; Gilroy, J.; Croy, R.R.D.; Boulter, D.
Nucleic Acids Res. 11, 2367-2380, 1983
A>Title: The vicilin gene family of pea (Pisum sativum L.): a complete cDNA coding sequence
A:Reference number: A93462; MUID:83220791; PMID:6687941
A:Accession: A03344
A:Molecule type: mRNA
```

```
A;Residues: 1-410 <LYC>
A;Cross-references: UNIPROT:P02854
A;Experimental source: cv. Feltham First, clones pDUB7 and pDUB4
A>Note: parts of this sequence, including the amino end of the mature protein, were deter
C;Comment: The gene that codes for this protein is part of a multigene family coding for
C;Superfamily: glycoprotein
C;Keywords: glycoprotein
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-410/Product: vicilin type B (fragment) #status predicted <MAT>
F;321-322/Cleavage site: Asn-Asp (unidentified proteinase) #status experimental
F;359/Binding site: carbohydrate (Asn) (covalent) #status experimental
```

Alignment Scores:

Pred. No.:	7,05e-67	Length:	410
Score:	1068.00	Matches:	209
Percent Similarity:	71.64%	Conservative:	84
Best Local Similarity:	51.10%	Mismatches:	104
Query Match:	29,98%	Indels:	12
DB:	1	Gaps:	5

US-10-728-323-1 (1-2032) x FWPMB (1-410)

```
QY 542 GTGAGGGAAGAAACATCTCGGAACAAACCCCTTTCTACTTCCCTCAAGGCGGTTTAGCAC 601
Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32
QY 602 CGCTACGGGAACAAACCGTAGGATCCGGGTCCTGAGAGTTTGACCAAGGTCAAGG 661
Db 33 LeuTyrlGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52
QY 662 CAGTTTCAGAACTCCAGNATCACGTATTGTCAGATCGAGCCCAAACTAACACTCTT 721
Db 53 IlePheGluAsnLeuGlnAsnTyrlArgLeuLeuGluTyrlLysSerLysProHisThrLeu 72
QY 722 GTTCTTCCCAAGCAGCTGATGCTGATACATCTCTGTTATCCAGCAAGCGCAAGCCACC 781
Db 73 PheLeuProGlnTyrlThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaThr 92
QY 782 GTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCAC 841
Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112
QY 842 AGAATCCCATCCGTTTCATTTCTTCTACATCTTGAACCGCATCAGACAGAACCTCAGA 901
Db 113 LysLeuProAlaGlySerIleAlaTyrlPheAlaAsnArgAspAsnGluGluProArg 132
QY 902 GTAGCTAAATCTCCATGCCCTTAAACACACCCCGCCAGTTTGAGGATTTCTTCCGCGG 961
Db 133 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152
QY 962 AGCAGCCGAGACCAATCATCTTCTGCAAGGCTTTCAGAGGAATACGTTGAGGCGCGC 1021
Db 153 GlyThrGlnAsnGlnLysSerSerLeuSerGlyPheSerLysAsnIleLeuGluAlaAla 172
QY 1022 TTCATGCGGAATTCATGAGATACGGAGGTCCTGTAGAGAGATATCAGAGAGTGTAG 1081
Db 173 PheAsnThrAsnTyrlGluGluIleGluLysValLeuLeuGluGlnGln-----Glu 189
QY 1082 CAAGAGGAGAGGCGCAGAGCGATCGGAGTACTCCGAGTAGTGAG---AACAAATGAAGGA 1138
Db 190 GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGluIleAsnGluLysAsn 209
QY 1139 GTGATAGTCAAAAGTGTCAAAGGAGCAGCGTTGAAGAACTTACTAAGCAGCGCTAAATCCGTC 1198
Db 210 ValIleValLysValSerArgAspGlnIleGluLeuSerLysAsnAlaLysSerSer 229
QY 1199 TCAAGAAAGGCTCCGAGAGAGAGGAGGATATCACAACCCCAATCAACTTGAGAGAGGC 1258
Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246
QY 1259 GAGCCCGATCTTCTTAAACAACTTTCGGAGGTTTATTTGAGGTGATGAGTCAAGCAGCAAGAGAAC 1318
Db 247 AsnProIleTyrlSerAsnLysPheGlyLysPhePheGluIleThrProGlu---LysAsn 265
```

```
QY 1319 CCCAGCTTCAGGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAAAGGAGCTTGG 1378
Db |||||
Db 266 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleAspValGlySerLeu 285
QY 1379 ATGCTCCACACTCAACTCAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGA 1438
Db |||||
Db 286 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 305
QY 1439 AACCTTGAACCTCGTCTGTAGAAAGAGACCAACACAGGGGGCGGCGGAGAGAG 1498
Db |||||
Db 306 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 321
QY 1499 GAGGACGAAGCAAGAGAGGAGGAGAACTAACACAGAGCTGCGTAGGTACACAGCGAGG 1558
Db |||||
Db 322 AspLysGluGluGluGlnGluGluThrSerLysGlnValGlnLeuTyrArgAlaLys 341
QY 1559 TTGAAGGAAGCGGATGTTTCATCATGCGCAGAGCTCATCCAGTACGCCATCAACGCTTCC 1618
Db |||||
Db 342 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 361
QY 1619 TCCGAACCTCACTCGCTTCGCTATCAACGCTGAAACAAACACACAGATCTTCCTT 1678
Db |||||
Db 362 SerAspLeuAsnLeuIleGlyLeuGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeu 381
QY 1679 GCAGGTGATAAGGACAATGTGTAGACCATGATAGAGAGCAAGCGAAGGATTTAGCATTC 1738
Db |||||
Db 382 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 401
QY 1739 CTGGGTCGGGTGAACAAGTTGAGAAG 1765
Db |||||
Db 402 ProGlySerSerHisGluValAspArg 410

RESULT 12
JQ2264
canavalin - jack bean
A:Species: Canavalia ensiformis (jack bean)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C:Accession: JQ2264; S19137
R:Ng, J.D.; Ko, T.P.; McPherson, A.
Plant Physiol. 101, 713-728, 1993
A:Title: Cloning, expression, and crystallization of jack bean (Canavalia ensiformis) ca
A:Reference number: JQ2264; MUID:94143475; PMID:8310055
A:Accession: JQ2264
A:Molecule type: mRNA
A:Residue: 1-445 <NGJ1>
A:Cross-references: UNIPROT:P50477; GB:X59467; GB:S76871; NID:g17976; PIDN:CAA42075.1; F
A:Experimental source: seed
R:Ng, J.D.; Stinchcombe, T.; Ko, T.P.; Alexander, E.; McPherson, A.
Plant Mol. Biol. 18, 147-149, 1992
A:Title: PCR cloning of the full-length cDNA for the seed protein canavalin from the jac
A:Reference number: S19137; MUID:92119225; PMID:1731967
A:Accession: S19137
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-445 <NGJ2>
A:Cross-references: EMBL:X59467; NID:g17976; PIDN:CAA42075.1; PID:g17977
C:Superfamily: glycinin
C:Keywords: storage protein

Alignment Scores:
Pred. No.: 1.74e-59 Length: 445
Score: 962.50 Matches: 198
Percent Similarity: 54.03% Conservative: 117
Best Local Similarity: 33.96% Mismatches: 101
Query Match: 27.02% Indels: 167
DB: 2 Gaps: 8

US-10-728-323-1 (1-2032) x JQ2264 (1-445)
```

```
QY 68 CCACGTGATCTTGTAGGATCCCTGTCTGGCTTCAGTTCTTCCACACCGCATGCCAAG 127
Db |||||
Db 8 ProLeuTrpLeuLeuGlyValValLeuLeuAlaSerValSerAlaSerPheAlaHis 27
```

```
QY 128 TCATCACCTTACCAGAGAGAAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
Db |||||
Db 28 Ser----- 28
QY 188 CAACAGGAAACCGATGACTTGAAGCAAAAGGATCGAGTCTCGTGCACCAAGCTCGAG 247
Db ----- 28
QY 248 TATGATCCTCGTTGTGTCTATCATCTCGAGGACACACTGGCACCACCACCAACCAAGTTCC 307
Db |||||
Db 29 -----GlyHisSerGly----- 32
QY 308 CCTCAGGGAGCGGACACACGTCGCCCAACCCGAGAGACTACGATGATGACCGCGCTCAA 367
Db |||||
Db 33 -----GlyGluAlaGluAspGluSerGluGlu 41
QY 368 CCCCAAGAGAGGAAGAGGAGCGGATGGGGACACAGCTGGACCGAGGGAGCGTGAAGAGAA 427
Db |||||
Db 42 SerArg----- 43
QY 428 GAAGACTGGAGACAACCAAGAGAGATTTGGAGGCGCACCAAGTCATCAGCAGCCCGGAAA 487
Db ----- 43
QY 488 ATAAGGCCGAAGGAAGAGAGAACACAGAGTGGGGAAACACACAGGTAGCATGTGAGG 547
Db ----- 43
QY 548 GAAGAAACATCTCGGAACAACCTTTCTACTTCCTCGTCAAGCGGCTTTAGCACCGCTAC 607
Db |||||
Db 44 -----AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe 60
QY 608 GGGACCAAAACGTTAGGATCGGGCTCTCGAGAGTTTGACCAAGGTCACAGGAGTTT 667
Db |||||
Db 61 LysAsnGlnHisGlySerLeuLeuGlnArgPheAsnGluAspThrGlnLysLeu 80
QY 668 CAGATCTCTCAGAACTACCGGTATTGTGCAGATCGAGGCGCAACCTTAACACTCTTCTTCTT 727
Db |||||
Db 81 GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu 100
QY 728 CCCAAGCAGCGTGTGTGATTAACATCTTGTATTCAGCAAGGCGCAAGCCACCGTAC 787
Db |||||
Db 101 ProHisHisSerAspSerAspLeuValLeuValLeuGluGlyGlnAlaIleVal 120
QY 788 GTAGCAATGCGCAATAACAGAAAGAGCTTAACTTGTGACGAGGCGCATCTCAGAGATC 847
Db |||||
Db 121 LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle 140
QY 848 CCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACACCAAGAACCTCAGAGTAGCT 907
Db |||||
Db 141 GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnAsnGlnAsnLeuArgIleLeu 160
QY 908 AAATCTCCATCGCCGTTAAACACCCCGCGGATTTGAGGATTTTCCCGCGGAGCAGC 967
Db |||||
Db 161 LysPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerThr 180
QY 968 CGAGACCAATCATCTACTTCGAGGCTTCAGAGGATACGTTGGAGGCGCGCTTCAAT 1027
Db |||||
Db 181 LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp 200
QY 1028 GCGGAATTCATATGACATACGAGGCTGTGTTAGAAGAGAATGACGAGGAGGTGACGAGAG 1087
Db |||||
Db 201 SerProTyrAspGluIleGluGlnThrLeuLeuGln----- 212
QY 1088 GAGAGGGCGAGCGGATGAGTACTCGGATAGTGAGAAACAATGAGAGGATGATGTC 1147
Db |||||
Db 213 -----GluGluGlnGluGlyValIleVal 220
QY 1148 AAAGTGTCAAAGGACGAGCTTGAAGAACTTACTAAGCACGCTAAATCCGTCCTCAAGAAA 1207
Db |||||
Db 221 LysMetProLysAspGlnIleGlnGluIleSerLysHisAlaGlnSerSerArgLys 240
```

```

QY 1208 GGCTCCGAAGAGGAGATATACCAACCCCAATCAACTTGAGAGAGCGGAGCCGAT 1267
Db 241 ThrLeuSerSerGln- - - - - ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1268 CTTTCTAACCACTTTGGAGTTATTTGAGTGAAGCCACAGCAAGAAAGAACCCCGAGCTT 1327
Db 258 TyrSerAsnAsnTyrGlyLeuTyrGluLeuThrProGlu- - - - - LysAsnSerGlnLeu 276
QY 1328 CAGGACCTGACATGATGCTCAGCTGTGTAGAGATCAAGAGAGAGCTTTGATGCTCCCA 1387
Db 277 ArgAspLeuAspIleLeuLeuAsnCysLeuGlnMetAsnGluGlyAlaLeuPheValPro 296
QY 1388 CACTTCAACTCAAGCCATGTTATCTGCTGCTCAACAAAGAACTGGAACCTTGAA 1447
Db 297 HisTyrAsnSerArgAlaThrValIleLeuValAlaAsnGluGlyArgAlaGluValGlu 316
QY 1448 CTGCTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGAGCGAA 1507
Db 317 LeuValGlyLeuGluGlnGlnGln- - - - - 326
QY 1508 GACGAAGAGAGAGGGA- - - - - AGTAACAGAGAGGTGGTAGGTACACAGCGAGTTGAAG 1564
Db 327 - - - - - -GlyLeuGluSerMetGlnLeuArgArgTyrAlaAlaThrLeuSer 341
QY 1565 GAAGCGATGCTTCATCATCGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA 1624
Db 342 GluGlyAspIleIleValIleProSerPheProValAlaLeuLysAlaSerAsp 361
QY 1625 CTCATCTGCTGCTGCTGCTATCAACGCTGAAACCAACACAGAAATCTTCTCTGAGGT 1684
Db 362 LeuAsnMetValGlyIleGlyValAsnAlaGluAsnGluArgAsnPheLeuAlaGly 381
QY 1685 GATAAGCAATGTGTAGACAGATAGAGAACCAAGCGAGGATTTAGCATTCCTCTGGG 1744
Db 382 HisLysGluAsnValIleArgGlnIleProArgGlnValSerAspLeuThrPheProGly 401
QY 1745 TCGGTTGACAACTGTGAGAGCTCATCAAAACAGAGAAATCTCATTGTGAGTGCT 1804
Db 402 SerGlyGluGluValGluGluLeuLeuGluAsnGlnLysGluSerTyrPheValAspGly 421
QY 1805 CGTCTCCAA 1813
Db 422 GlnProArg 424

RESULT 13
S00281
C:Species: Canavalia gladiata (sword bean)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: S00281; S04598
R:Yamauchi, D.; Nakamura, K.; Asahi, T.; Minamikawa, T.
Eur. J. Biochem. 170, 515-520, 1988
A:Title: cDNAs for canavalin and concanavalin A from Canavalia gladiata seeds. Nucleotide
s.
A:Reference number: S00281; MUID:88111636; PMID:3338449
A:Accession: S00281
A:Molecule type: mRNA
A:Residues: 1-445 <YAM>
A:Cross-references: UNIPROT:P10562; EMBL:X06733; NID:g18003; PIDN:CAA29910.1; PID:g18004
R:Takei, Y.; Yamauchi, D.; Minamikawa, T.
Nucleic Acids Res. 17, 4381, 1989
A:Title: Nucleotide sequence of the canavalin gene from Canavalia gladiata seeds.
A:Reference number: S04598; MUID:89296493; PMID:2740227
A:Accession: S04598
A:Molecule type: DNA
A:Residues: 1-160, 'K', 162-445 <YAK>
A:Cross-references: EMBL:X15076; NID:g18006; PIDN:CAA333172.1; PID:g18007
C:Genetics:
A:Introns: 115/1; 173/3; 200/3; 290/3; 381/1
C:Superfamily: glycinin
C:Keywords: seed

Alignment Scores:

```

```

Pred. No.: 2,82e-59 Length: 445
Score: 99.50 Matches: 197
Percent Similarity: 54.03% Conservative: 118
Best Local Similarity: 33.79% Mismatches: 101
Query Match: 26.94% Indels: 167
DB: 2 Gaps: 8

US-10-728-323-1 (1-2032) x S00281 (1-445)

QY 68 CCACATGATCTGTGTCTAGGATCTCTGTCTGGTCTCAGTTTCTGCAACGATGCAAG 127
Db 8 ProLeuTrrpLeuLeuLeuGlyValLeuLeuAlaSerValSerPheAlaHis 27
QY 128 TCATCACTTACCAGAGAAACACAGAACCCCTGCGCCACAGAGTGCCTCCAGAGTTGT 187
Db 28 Ser- - - - - 28
QY 188 CAACAGAACCGATGACTTGAAGCAAAAGGATCGAGTCTCGTGTCCACCAAGCTCGAG 247
Db 28 - - - - - 28
QY 248 TATGATCTCTGTGTGTCTATGATCTCTGAGGACACACTGGCACCAACCAACCAACGTTCC 307
Db 29 - - - - - -GlyHisSerGly- - - - - 32
QY 308 CTTCCAGGGAGCGGACACGCTGCGCCCAACCCGAGAGACTAGCATGATACCGCCGTCAA 367
Db 33 - - - - - -GlyGluAlaGluAspGluSerGluGlu 41
QY 368 CCCCAGAGAGAGAGGAGGCGGATGGGACACAGCTGGACCGGAGGAGCGTGAAAGAGAA 427
Db 42 SerArg- - - - - 43
QY 428 GAAGACTGGAGACAACCAAGAGAGATTGGAGGCGCAACCAAGTCTATCAGCAGCCACGAA 487
Db 43 - - - - - 43
QY 488 ATAAGCCCGAAGGAGAGAGAGAGAACAAAGTGGGGAGAACCAAGGTAGCCATGTGAGG 547
Db 43 - - - - - 43
QY 548 GAAGAAACATCTCGGAACAACCTTTCTACTTCCGTCAAGCGGTTTAGCACCCGCTAC 607
Db 44 - - - - - AlaGlnAsnAsnProTyrLeuPheArgSerAsnLysPheLeuThrLeuPhe 60
QY 608 GGGAAACCAACCGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCAAGGCGAGTTT 667
Db 61 LysAsnGlnHisGlySerLeuArgLeuLeuGlnArgPheAsnGluAspThrGluLysLeu 80
QY 668 CAGATCTCCGAATCACCGTATTGTGAGATCGAGGCGCAACCTTAACACTCTTGTCTTT 727
Db 81 GluAsnLeuArgAspTyrArgValLeuGluTyrCysSerLysProAsnThrLeuLeuLeu 100
QY 728 CCCAAGCAGCTGATGCTGATAAATCCTTGTATTCTCAGCAAGGGCAAGCCACCGTGACC 787
Db 101 ProHisHisSerAspSerAspLeuValLeuValLeuGluGlyGlnAlaIleLeuVal 120
QY 788 GTAGCAATGGCAATAAACAGAAAGAGCTTTAATCTTTGACGAGGCGCATGCACTCAGAATC 847
Db 121 LeuValAsnProAspGlyArgAspThrTyrLysLeuAspGlnGlyAspAlaIleLysIle 140
QY 848 CCATCCGTTTCATTTCTTACATCTTTGAACCCGCATGACCAACAGAACCTCAGAGTAGCT 907
Db 141 GlnAlaGlyThrProPheTyrLeuIleAsnProAspAsnGlnAsnLeuArgIleLeu 160
QY 908 AAAATCTCATGCCGTTAAACACACCCGCGCTTTGAGGATTTCTTCCCGGCGGACGAC 967
Db 161 AsnPheAlaIleThrPheArgArgProGlyThrValGluAspPhePheLeuSerSerThr 180
QY 968 CGAGACCAATCATCTTACTTTCAGGCGCTTCAGCAGGAATACGTTGGAGCGCCGCTTCAAT 1027
Db 181 LysArgLeuProSerTyrLeuSerAlaPheSerLysAsnPheLeuGluAlaSerTyrAsp 200

```


Search completed: August 24, 2005, 10:06:34
Job time : 54.4416 secs

```
QY 806 AGAAGAGCTTTAACTTCGACGAGGCCATGCACCTCAGAATCCCATCCGTTTCATTTCC 865
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 866 TACATCTTGAACCGCCATGACAAACCACTCAGAGTAGCTAAATCTCCATGCCCGTT 925
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 881 TyrLeuAlaAsnGlnAspAsnArgGluLysLeuThrIleAlaValLeuHisArgProVal 300
QY 926 AACACACCGCGCCAGTTTCGAGGATTTCTCCGGCGAGCAGCCGAGACCAATCATCCTAC 985
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 931 AsnAsnProGlyGlnPheGlnLysPhePheProAlaGlyGlnGluAsnProGlnSerfyr 320
QY 986 TTGACGGCTTCAGCAGGAATACGTTGGAGCGCCCTTCAATCCGGAATTCATCAGATA 1045
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1046 CGGAGGGTCTGTTAGAAGAGATTCAGAGGTGAGCAAGAGAGAGAGGGCAGAGCGGA 1105
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 336 ArgSerGluGlnLeuAspGluLeuProGlyGlyArgGlnSerHisArgArgGlnGln--- 354
QY 1106 TGGAGTACTCGGAGTAGTCAGAACATGAAGAGAGTAGTCAAGTGTCAAAGGAGCAC 1165
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 355 -----GlyGlnGlyMetPheArgLysAlaSerGlnGln 366
QY 1166 GTTGAAGAACTTACTAAGCAGCTAAATCCGTCTCAAAGAAAGGCTCCGAGAGAGAGGA 1225
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 367 IleArgAlaLeuSerGlnGlyAlaThrSerProArgGlyLysGlySerGluGlyTyrAla 386
QY 1226 GATATCACCACCAATCAACTTGAGAGAGGGAGCGCCGATCTTTCTAACTTTGGG 1285
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 387 -----PheAsnLeuLeuSerGlnThrProArgTyrSerAsnGlnAsnGly 401
QY 1286 AAGTTATTGAGGTGAAGCGACAGACAGAACCCCGCTTCAGGACCTGGACATGATG 1345
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 402 ArgPheTyrGluAlaCysProArgAsnPheGlnGlnGlnLeuArgGluValAspSerSer 421
QY 1346 CTCACCTGTGTAGAGATCAAAGAGAGGCTTTGATGCTCCACACTTCAACTCAAAGGCC 1405
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 422 ValValAlaPheGluIleAsnLysGlySerIlePheValProHisTyrAsnSerLysAla 441
QY 1406 ATGGTTATCGTGTGCTCAACAAAGAACTGAAACCTTGAACCTGTGGCT-----GTA 1459
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 442 ThrPheValValLeuValThrGluGlyAsnGlyHisValGluMetValCysProHisLeu 461
QY 1460 AGAAAGAGCAACACAGAGGGGACGGCGGGAAGAGAGGAGGAGACGACGACGAAGAGAG 1519
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 462 SerArgGlnSerSerAspTrpSerSerArgGluGluGluGluGluGluGluVal 481
QY 1520 GAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCGAGGTGAAGGAAGCGATGTGTTTC 1579
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 482 GluArgArgSerGlyGlnTyrLysArgValArgAlaGlnLeuSerThrGlyAsnLeuPhe 501
QY 1580 ATCATCCGACGAGCTCATCGTAGCCATCAAGCTTCC-----TCCGAACCTCCATCTG 1633
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 502 ValValProAlaGlyHisProValThrPheValAlaSerGlnAsnGluAspLeuGlyLeu 521
QY 1634 CTTGGCTTCGGTATC---AACCGTGAACAAACCAACAGAACTCTCTTCGCGGTGATAG 1690
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 522 LeuGlyPheGlyLeuTyrAsnGlyGlnAspAsnLysArgIlePheValAlaGly---Lys 540
QY 1691 GACAATGTGATAGACCATAGAGAGCAAGCAAGGATTTAGCATTCCTCGGGTCGGGT 1750
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 541 ThrAsnAsnValArgGlnTrpAspArgGlnAlaLysGluLeuAlaPheGlyValGluSer 560
QY 1751 GAACAAGTTGAGAAGCTCATCAAAACACAG---AAGGAATCTCACTTTGTGAGTCTGCT 1807
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 561 ArgLeuValAspGluValPheAsnAsnAsnProGlnGluSerTyrPheValSerGlyArg 580
QY 1808 -----CCTCAATCTCAATCTCAATCTCGCTGCTCTCT 1840
Db ::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 581 AspArgArgGlyPheAspGluArgArgGlySerAsnAsnProLeuSerPro 597
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 124.355 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3562

Sequence: 1 aataacatataattcatc.....cgtttggtggtttcttcc 2032

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model
-Q=/cgn2 1/USPTO.spool/US10728323/runat 23082005 124355 29210/app query.fasta 1.4757
-DB=Uniprot 03 -QFMT=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pept -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1 291 @runat 23082005 124355 29210 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3286	92.3	626	1 AL12 ARAHY	P43238 arachis hyp
2	3041	85.4	614	1 AL11 ARAHY	P43237 arachis hyp
3	2887.5	81.1	580	2 Q6PSU3	O6psu3 arachis hyp
4	2172.5	61.0	428	2 Q6PSU4	O6psu4 arachis hyp
5	1495.5	42.0	299	2 Q6PSU5	O6psu5 arachis hyp
6	1435	40.3	303	2 Q6PSU6	O6psu6 arachis hyp
7	1302	36.6	621	2 Q6EB1	O7xtt2 glycine max
8	1290	36.2	533	2 Q6EB2	O6ebc1 lupinus alb
9	1284	36.0	621	2 Q94Y0	O948y0 glycine max
10	1282	36.0	605	2 Q94X2	O941x2 glycine max
11	1281	36.0	605	1 GLCA SOYBN	P13916 glycine max
12	1267	35.6	571	1 CVCA PEA	P13915 pisum sativ
13	1265.5	35.5	623	2 Q948X9	O948x9 glycine max
14	1258.5	35.3	545	2 Q41674	O41674 vicia narbo
15	1251	35.1	613	2 Q9M3X6	O9m3x6 pisum sativ
16	1251	35.1	639	1 GLCA SOYBN	P11827 glycine max

17	1190	33.4	559	2 Q9FZP9	Q9fzp9 glycine max
18	1183.5	33.2	418	2 Q840I1	Q840i1 lens culina
19	1180.5	33.1	543	2 Q22120	Q22120 glycine max
20	1178.5	33.1	459	1 VCLC PEA	P13918 pisum sativ
21	1168	32.8	415	2 Q84U10	Q84u10 lens culina
22	1165	32.7	416	2 Q22121	Q22121 glycine max
23	1156	32.5	439	1 GLCB SOYBN	P25974 glycine max
24	1152	32.3	439	2 Q93VL9	Q93vl9 glycine max
25	1148	32.2	438	2 Q43626	Q43626 pisum sativ
26	1145	32.1	415	2 Q702P1	Q702p1 pisum sativ
27	1135	31.9	415	2 Q702P0	Q702p0 pisum sativ
28	1123.5	31.5	463	2 Q41677	Q41677 vicia narbo
29	1108.5	31.1	463	1 VCL VICFA	P08438 vicia faba
30	1082	30.4	396	2 Q84UB3	Q84ub3 glycine max
31	1068	30.0	410	1 VCLB PEA	P02854 pisum sativ
32	1030	28.9	518	2 Q9M3X8	Q9m3x8 lens culina
33	962.5	27.0	445	1 CANA CANEN	P50477 canavalia e
34	959.5	26.9	445	1 CANA CANGL	P10562 canavalia g
35	910	25.5	593	2 Q9SEW4	Q9sew4 juglans reg
36	907.5	25.5	810	2 Q9ZWI3	Q9zwi3 cucurbita m
37	880	24.7	481	2 Q7YIC1	Q7yic1 juglans nig
38	865.5	24.3	666	2 Q9SPL4	Q9spl4 macadamia i
39	863.5	24.2	625	2 Q9SPL3	Q9spl3 macadamia i
40	849.5	23.8	666	2 Q9SPL5	Q9spl5 macadamia i
41	839	23.6	605	1 VCLA GOSHI	P09799 gossypium h
42	825.5	23.2	436	1 PHSA PHAVU	P07219 phaseolus v
43	806.5	22.6	430	2 Q41115	Q41115 phaseolus v
44	805.5	22.6	430	2 Q43633	Q43633 phaseolus v
45	804	22.6	423	1 PHG2 PHALU	Q43617 phaseolus l

ALIGNMENTS

RESULT 1

ID	AL12 ARAHY	STANDARD;	PRT;	626 AA.
AC	P43238;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Allergen Ara h 1, clone P418 precursor (Ara h 1).			
OS	Arachis hypogaea (peanut).			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;			
OC	Arachis.			
OX	NCBI_TaxID=3818;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Florunner;			
RX	MEDLINE=96013631; PubMed=7560062;			
RA	Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;			
RT	"Recombinant peanut allergen Ara h I expression and IGE binding in			
RT	patients with peanut hypersensitivity.";			
RL	J. Clin. Invest. 96:1715-1721(1995).			
RN	[2]			
RP	CARBOHYDRATE-LINKAGE SITE ASN-521.			
RX	MEDLINE=20455243; PubMed=1098264; DOI=10.1006/abio.2000.4737;			
RA	Kolarich D., Altmann F.;			
RT	"N-glycan analysis by matrix-assisted laser desorption/ionization mass			
RT	spectrometry of electrophoretically separated nonmammalian proteins:			
RT	application to peanut allergen Ara h I and olive pollen allergen Ole e			
RL	Anal. Biochem. 285:64-75(2000).			
CC	-!- ALLERGEN: Causes an allergic reaction in human.			
CC	-!- SIMILARITY: Belongs to the 7S seed storage protein family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			

or send an email to license@isb-sib.ch.

CC EMBL; L34402; AAB00861.1; -;
CC HSSP; P25974; 1LRJ
DR InterPro; IPR06045; Cupin.
DR InterPro; IPR00713; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 626 Allergen Ara h 1, clone P41B.
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 626 AA; 71345 MW; 1A6BBB411490D0E3 CRC64;

Alignment Scores:

Pred. No.: 2 34e-207 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 1 Gaps: 0

US-10-728-323-1 (1-2032) x AL12_ABAHY (1-626)

QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGGATCCTGTCTCGTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCGCAACGCATGCCAAGTCATCACCCTTACCAGAGAAAACAGAACCCCTCGCCCGAG	169
DB	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTCAAGCAAAAGCATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCGAGTATGATCTCGTGTGCTATGATCCTCGAGGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluLysArgProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACAGTGGCGCCCAACCCGGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCGCGCTCAACCCCAAGAGAGGAGCGCGATGGGGACAGCTGGACCG	409
DB	101	AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro	120
QY	410	AGGAGCGTGAAGAAGAAGACTGGAGACCAACCAAGAGAAGATTGGAGGCGCAAGT	469
DB	121	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer	140
QY	470	CATCAGCACCCGGAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnLysArgPhe	160
QY	530	CGAGTAGCATGTGAGGAAGAAACATCTCGGAACACCCCTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCGAGAGGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAGGTCAGGCAGTTTCAGATCTCCAGATCACCGTATTGTGCAGATCCAGGCCAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTCTTCCCAAGCAGCTGATGCTGATAACATCTCTGTTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln	240
QY	770	GGCAAGCCCGCTGACCGTAGCAATAGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829

DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACCTCAGAAATCCCATCCGTTTCATTTCTACATCTTCAACCGCCATCACAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCCGCCAGTTTCAAGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGGAGCAGCGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGGCGGCTTCAATGCCGAATTCATGAGATACGGAGGGTCTGTTAGACAGAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn	340
QY	1070	GCAGGAGGTGACAAAGAGAGAGAGCGGATGAGTACTCGAGTAGTAGTACAGAAC	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn	360
QY	1130	AATGAAGGAGTATAGTCAAAAGTGTCAAAGGAGCAGTTGAAGAACTTAAAGCACGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCATCACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGAGCCGATCTTTAAACAACTTTGGGAAGTTATTGAGTGAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAAGACCCAGCTTCAGACCTGCACATGATGCTCAGCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu	440
QY	1370	CGAGCTTTGATGCTCCACACTCAACTCAAGGCCATGTTATCGTCTGCTCAACAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACTTTGAACTCTGCTGCTTAAGAAAGAGCAACACAGAGGGGAGCGGG	1489
DB	461	GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGTTCAAGAGGAGCGATGTTTCATGCCAGCAGCTCATCCAGTACCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACCTCATCTGCTTGGCTTCAACGCTCAACGCTGAAACCAACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg	540
QY	1670	ATCTTCTTGAGGTGATAAGACAAATGTGATAGACCAGATAGAGAACCAAGCGAGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTGGGTGGGTGAACTTCAGAGCTCATCAAAACACAGAGGAATCT	1789
DB	561	LeuAlaPheProLysLysGlyGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTTGAGTGTCTGCTCCTCAATCTCAATCTCGCTCGCTCTCTGAGAGAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGlyLysGlu	600
QY	1850	TCTCTGAGAAAGAGGATCAAGAGGAGAAACCAAGAGGAGGAGGCTCTCTCTTCA	1909

Db	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGCGTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 2			
AL11	ARAHY	STANDARD;	PRT; 614 AA.
AC	P43237;		
DT	01-NOV-1995	(Rel. 32, Created)	
DT	01-NOV-1995	(Rel. 32, Last sequence update)	
DT	10-OCT-2003	(Rel. 42, Last annotation update)	
DE	Allergen Ara h 1, clone P17 precursor (Ara h I).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Florunner;		
RX	MEDLINE=96013631; PubMed=7560062;		
RA	Burke A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;		
RA	"Recombinant peanut allergen Ara h I expression and IgE binding in		
RT	patients with peanut hypersensitivity.";		
RL	J. Clin. Invest. 96:1715-1721 (1995).		
RN	[2]		
RP	CARBOHYDRATE-LINKAGE SITE ASN-516.		
RX	MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;		
RA	Kolarich D., Altmann F.;		
RT	"N-glycan analysis by matrix-assisted laser desorption/ionization mass		
RT	spectrometry of electrophoretically separated nonmammalian proteins:		
RT	application to peanut allergen Ara h 1 and olive pollen allergen Ole e		
RT	1.";		
RL	Anal. Biochem. 285:64-75 (2000).		
CC	-!- ALLERGEN: Causes an allergic reaction in human.		
CC	-!- SIMILARITY: Belongs to the 7S seed storage protein family.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; L38853; AAA60336.1; --		
DR	HSSP; P25974; 11PJ.		
DR	GlycoSuiteDB; P43237; --		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin region.		
DR	Pfam; PF00190; Cupin; 2.		
KW	Allergen; Glycoprotein; Signal.		
FT	SIGNAL 1 25 Potential.		
FT	CHAIN 26 614 Allergen Ara h 1, clone P17.		
FT	CARBOHYD 516 516 N-linked (GlcNAc...).		
FT	FTId-CAR 000218.		
SQ	SEQUENCE 614 AA; 70283 MW; 1DDACF217EBC5F31 CRC64;		
Alignment Scores:			
Pred. No.:	2.9e-191	Length:	614
Score:	3041.00	Matches:	595
Percent Similarity:	96.03%	Conservative:	9
Best Local Similarity:	94.59%	Mismatches:	7
Query Match:	85.37%	Indels:	18
DB:	1	Gaps:	7
US-10-728-323-1 (1-2032) x AL11_ARAHY (1-614)			
QY	50	ATCAGAGGGGGTTTCTCCCATGATGCTGTGCTAGGATCCTTGTCTCGCTTCAGTT	109

Db	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAAACAGAGAACCCCTGCGCCAG	169
Db	21	SerAlaThrGlnAlaLys--SerProTyr--ArgLysThrGluAsnProCysAlaGln	38
QY	170	AGTGCTTCAGAGTTGTCACAGGAAACCGGATGACTTGAAGCAAAAGGATCGAGTCT	229
Db	39	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	58
QY	230	CGCTGCACCAAGCTCGAGTATGATCTCGTGTGTCTATGATCTCTCGAGGACACTGGC	289
Db	59	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly	74
QY	290	ACCACCAACCAAGTTTCCCTCCAGGGAGCGGACACAGTGGCGGCCCAACCCGGAGACTAC	349
Db	75	AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	94
QY	350	GATGATGACCGCGCTCAACCCCGAAGAGAGAGAGGCGCGATGGGGACAGCTGGACCG	409
Db	95	AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGluPro	114
QY	410	AGGGAGCGTGAAGAGAGAAAGACTGGAGACAACCAAGAGAGAAATTGGAGGCGACCAAGT	469
Db	115	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer	134
QY	470	CATCAGACGCCAGGAAATAAGGCCCGAAGAGAGAGAGAGAAACAGAGTGGGGAACA	529
Db	135	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnTrpGlyThr	154
QY	530	CCAGTAGCCATGTGAGGAGAAACATCTCGAACAACCCCTTCTACTCCGCTCAAGG	589
Db	155	ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	174
QY	590	CGTTTAGCACCGCTACGGGAAACCAAAACGCTAGGATCCGGCTCTCGAGAGTTTGAC	649
Db	175	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	194
QY	650	CAAAGTCAAGGCAGTTTCCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAA	709
Db	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
QY	710	CCTAACACTCTTGTCTTCCCAAGCAGCTGATGTGATGAACATCTCTTGTATTCAGCAA	769
Db	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
QY	770	GGCAAGCCACCGTCAGCGTAGCAATGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
QY	830	GGCCATGCACCTCAGAAATCCCATCCGGTTTTCATTTCTTCTACATCTTGAACCGCCATGCAAC	889
Db	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
QY	890	CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCGCAGTTTAGGAT	949
Db	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTCTTCCCGCGAGCAGCGGAGACCAATCATCTACTTTCAGGGCTTCAGAGGATACG	1009
Db	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTGGAGGCGCGCTTCAATTCGGAATTCATATGAGATACGGAGGGTGTGTGTAGAAGAAAT	1069
Db	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	334
QY	1070	GCAGAGGTGACCAAGAGAGAGAGGCGGAGCGCATGGAGTACTCGGAGTAGTAGAAC	1129
Db	335	AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
QY	1130	AATGAGGAGTAGTAGTCAAGTGTCAAGGAGCAGTTGAAGACTTACTTAAGCAGCT	1189

Db 354 AsnGluGlyValrIleVallysValSerLysGluHisValGlnGlnLeuThrLysHisAla 373

QY 1190 AAATCGTCTCAAGAAAGCTCCGAAGAGAGGAGATATCACCAACCAATCAACTTG 1249

Db 374 LysSerValSerLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392

QY 1250 AGAGAAGCGAGCCGATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309

Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412

QY 1310 AAGAAGAACCCAGCTTCAGACCTGGAGATATGCTCACCTGTGTAGATCAAGAA 1369

Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu 432

QY 1370 GGAGCTTTGATGTCCTCCACACTCAACTCAAGGCCATGCTTATCGTCTGCTCAACAA 1429

Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452

QY 1430 GGAATCTGGAACCTTGAATCTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGCGCG 1489

Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 472

QY 1490 GAA-----GAAGAGAGACGACGAGACGAGAGAGAGAGAGAGAGAGAGAGAG 1540

Db 473 GluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGluVal 492

QY 1541 CGTAGGTACACAGCGAGGTGAAGAGCGGATGTTTCATCATGCCAGCCTCATCCA 1600

Db 493 ArgArgTyThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512

QY 1601 GTAGCATCAACCGTCTCCGAATCTCATCTGCTTGGCTTCGGTATCAACGCTCAAAAC 1660

Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532

QY 1661 AACCCAGAGATCTCTTCAGGTATAGAGCAATGTATGATACAGCAGATAGAGCA 1720

Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552

QY 1721 GCGAAGGATTTAGCATCTCCCTGGGTCCGAGTGAAGTGAAGTCAATCAATAACCAAG 1780

Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGln 572

QY 1781 AAGGAATCTCACTTGTGTAGTCTCGTCTCTCAATCTCAATCTCAATCTCGTCTCT 1840

Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588

QY 1841 GAGAAAGAGTCTCCTGAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGTCCA 1900

Db 589 -----SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605

QY 1901 CTCCTTTCAATTTTGAAGCTTTTAAAC 1927

Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 3

Q6PSU3 PRELIMINARY; PRT; 580 AA.

AC Q6PSU3;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Conarachin (Fragment).

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC euroside I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RN [1]

RP SEQUENCE FROM N.A.

RA Li H., Wang L., Liao B., Yan Y., Lin X., Huang S.;

RL Submitted (MAR-2004) to the EMBL/GenBank/DBSJ databases.

DR EMBL; AY581852; AAT00597.1; --

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR007113; Cupin region.

DR InterPro; IPR011051; RmlC_Like_cupin.

PFam; PF00190; Cupin; 2.

FT NON_TER 580

SQ SEQUENCE 580 AA; 66575 MW; 247AA144CCC1F36D CRC64;

Alignment Scores:

Score: 3,478-181 Length: 580

Pred. No.: 2887.50 Matches: 562

Percent Similarity: 97.11% Conservative: 9

Best Local Similarity: 95.58% Mismatches: 6

Query Match: 81.06% Indels: 11

DB: 2 Gaps: 6

US-10-728-323-1 (1-2032) x Q6PSU3 (1-580)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGGATCTTGTCTCTGCTTCAGTT 109

Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGCATGCCAAGTCATCACTTACCAAGAAAGAAACAGAGAAACCCCTGCCCCAG 169

Db 21 SerAlaThrGlnAlaLys---SerProTyrr---ArgLysThrGluAsnProCysAlaGln 38

QY 170 AGTGCCTCCAGAGTTGTCAACAGGAACCGGATGATTTGAAGCAAAAGGCATGCCAGTCT 229

Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58

QY 230 CGTGCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTCGAGGACACATCGC 289

Db 59 ArgCysThrLysLeuGluTyrrAspProArgCysValTyrrAsp-----ThrGly 74

QY 230 ACCACCAACCAACGTTCCCTCCAGGGAGGCGACACGTCGCCGCCAACCCCGAGACTAC 349

Db 75 AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyrr 94

QY 350 GATGATGACCCCGTCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409

Db 95 AspAspAspArgGlnProArgGlnGluGlyArgGlyArgGlyArgGlyArgGlyArgGly 114

QY 410 AGGGAGCGTGAAGAGAAAGAGACTGGAGACCAACCAAGAGAGAGAGAGAGAGAGAGAG 469

Db 115 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134

QY 470 CATCAGAGCCACGAGAAATTAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529

Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGlyArgGlyArgGlyArgGlyArg 154

QY 530 CCAGTAGCCATGTGAGGAGAGAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGG 589

Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrrPheProSerArg 174

QY 590 CGTTTATGACCCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCTCAGAGGTTTCAAC 649

Db 175 ArgPheSerThrArgTyrrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194

QY 650 CAAGGTCAAGGCAGTTTCCAGATCTCCAGATCAACCGTATTTGTCAGATCGAGGCCAAA 709

Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214

QY 710 CCTAACACTCTTGTCTTCTCCCAAGCAGCTGATGCTGATATACTCTTGTATTCAGCAA 769

Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234

QY 770 GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGACG 829

Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254

QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 889

Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrrIleLeuAsnArgHisAspAsn 274

QY	1304	CCAGACAAGAAAGACCCCACTTCAGGACCTGGACATGATGCTCACTCGTGTAGAGATC	1363
Db	221	ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIle	240
QY	1364	AAAGAAAGAGCTTTGATGCTCCCACTTCACACTTCAACTCAAAAGGCCATGGTTATCGTCGTGC	1423
Db	241	LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal	260
QY	1424	AACAAAGAACTGGAAACCTTGAACCTCTGGCTCTTAAGAAAGAGCAACAACACAGAGGGGA	1483
Db	261	AsnLysGlyThrGlyAsnLeuGluLeuValAlaValargLysGluGlnGlnArgGly	280
QY	1484	CGCGGGGAAGAGAGGAGGACCAAGACCAAGAGAGGAGGAAAGTAAACAGAGAGGTCCGT	1543
Db	281	ArgArgGluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArg	300
QY	1544	AGGTACACAGGAGGTTGAAGGAAGGCGATGTGTTTCATATGCCAGCAGCTCATCCAGTA	1603
Db	301	ArgTyrThrAlaArgLeuLysGluLysPheIleMetProAlaAlaHisProVal	320
QY	1604	GCCATCAACGCTTCTCCGAACCTCCATCTGCTTCGCTTCGGTATCAACGCTGAAAAACAAC	1663
Db	321	AlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn	340
QY	1664	CACAGAATCTTCCTTGCAGGTGATAAGGACAATGTATAGACCAAGATAGAGAAGCAAGCG	1723
Db	341	HisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAla	360
QY	1724	AAGATTATGACATTCCTCGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAAAG	1783
Db	361	LysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLys	380
QY	1784	GAATCTCACTTTGTGAGTGCCTCAATCTCAATCTCAATCTCGCTCGTCTCTCTGAG	1843
Db	381	GluSerHisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlu	400
QY	1844	AAAGAGTCTCTCGAAGAGAGGATCAAGAGGAGGAGGAAAAACCAAGAGGAGGAGGGTCCACTC	1903
Db	401	LysGluSerProGluLysGluAspGlnGluGluGlnGlyGlyLysGlyProLeu	420
QY	1904	CTTTCAATTTGAAGGCTTTTAAAC	1927
Db	421	LeuSerIleLeuLysAlaPheAsn	428
RESULT 5			
Q6PSU5	PRELIMINARY; PRT; 299 AA.		
ID	Q6PSU5		
AC	Q6PSU5;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Conarachin (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY581850; AAT00595.1; -.		
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin region.		
DR	InterPro; IPR011051; RmlC_Like_cupin.		
DR	Pfam; PF00190; Cupin; 1.		
FT	NON TER		
SQ	SEQUENCE 299 AA; 33604 MW; 594C3AB0C2FD49EB CRC64;		
Alignment Scores:			
Pred. No.:	8.28e-90	Length:	299

```

Q6PSU6
ID Q6PSU6 PRELIMINARY; PRT; 303 AA.
AC Q6PSU6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Conarachin (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang L., Liao B., Li H., Yan Y., Lin X., Huang S.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581849; AAT00594.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin_region.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 1.
FT NON TER 1
SQ SEQUENCE 303 AA; 34133 MW; 5B4E21BEB48654EC CRC64;

Alignment Scores:
Pred. No.: 7.84e-86 Length: 303
Score: 1435.00 Matches: 291
Percent Similarity: 95.50% Conservative: 6
Best Local Similarity: 93.57% Mismatches: 2
Query Match: 40.29% Indels: 12
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x Q6PSU6 (1-303)
QY 1004 AATACGTGGAGGCGCCCTTCATGCGGANTTCAATCAGATACGGAGGTGCTGTAGAA 1063
DB 2 AsnThrLeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGlu 21
QY 1064 GAGATGTCAGAGGTGAGCAAGAGGAGAGAGGCGAGCGATGCTACCGAGTAGT 1123
DB 22 GluAsnAlaGlyGlyGluGluGluValGlyGlnArgArgSerThrArgSerSer 41
QY 1124 GAGAACAAATGAGAGGTATGATCAAGTGTCAAAGAGACACCTTGAAGAACTTACTAG 1183
DB 42 Asp---AsnGluGlyValIleValIleValIleValIleValIleValIleValIle 60
QY 1184 CACGCTAAATCCCTCTCAAGAAAGGCTCCGAGAGAGGAGGAGATATCACCAACCAATC 1243
DB 61 HisAlaLysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIle 79
QY 1244 AACTTGAGAGAGCGAGCGCCGATCTTCTAACAACTTTGGGAGATTATTGAGGTGAAG 1303
DB 80 AsnLeuArgAspGlyGluProAspLeuSerAsnAsnAsnPheGlyArgLeuPheGluValLys 99
QY 1304 CCAGACAAAGAAACCCCGCTTACAGACCTCGACATGATCTACCTGTGTGATGATC 1363
DB 100 ProAspLysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIle 119
QY 1364 AAGAAGAGCTTTGATGCTCCACACATCTCAATCAAGCCCATGTTATCTGCTCGTC 1423
DB 120 LysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValVal 139
QY 1424 AACAAAGGAATCGAAACCTTGAACTCGTGGCTGTGAAGAAAGACAAACAGAGGGGA 1483
DB 140 AsnLysGlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGly 159
QY 1484 CGCGCGGAA-----GAAGAGGAGGACCAAGACCAAGAGAGGAGGAAGTAACAGA 1534
DB 160 ArgArgGluGlnGluTrpGluGluGluGluAspGluGluGluGlySerAsnArg 179
QY 1535 GAGGTGCTAGGTACACAGAGCGAGGTGAAGAGGCGATGTTCATCATGCCAGCAGCT 1594

180 GluValArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAla 199
1595 CATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTCTGCTGGCTTCGGTATCAACGCT 1654
200 HisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAla 219
1655 GAAACCAACACACAGATCTCTTCAGAGTGATAGGACAATGTGTATACACAGATAGAG 1714
220 GluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGlu 239
1715 AAGCAAGCGAAGATTAGCATCTCCCTGGGTGCGGTGAACAAGATTGAGAAGCTCATCAA 1774
240 LysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLys 259
1775 AACCAGAGGAATCTCATTCTGTGAGTCTCTCTCAATCTCAATCTCAATCTCCGTCG 1834
260 AsnGlnArgLysHisPheValSerAlaArgProGlnSerGlnSerProSer----- 277
1835 TCTCCTGAGAAAGAGTCTCTCAGAAAGAGGATCAAGAGAGGAGAAACCAAGAGGAG 1894
278 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLys 292
1895 GGTCCACTCTTTCAATTTTGAAGGCTTTTAAC 1927
293 GlyProLeuLeuSerIleLeuLysAlaPheAsn 303

RESULT 7
Q7XXT2 PRELIMINARY; PRT; 621 AA.
AC Q7XXT2;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Prepro beta-conglycinin alpha prime subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RA Urade R., Nakatani H., Nakano C.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB113351; BAC78524.1; -.
DR PDB; 1UIK; X-ray; A/B/Cs-.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 621 AA; 72247 MW; 0DF4B6E33737B7F CRC64;

Alignment Scores:
Pred. No.: 5.1e-77 Length: 621
Score: 1302.00 Matches: 286
Percent Similarity: 61.02% Conservative: 121
Best Local Similarity: 42.88% Mismatches: 170
Query Match: 36.55% Indels: 90
DB: 2 Gaps: 19

US-10-728-323-1 (1-2032) x Q7XXT2 (1-621)
QY 50 ATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 109
DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
QY 110 TCTGCAACCCATGCCAGTCACTACCTTACCAAGAGAAACAGAGAACCCCTGCGGCCAG 169
DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProSerHisAsn 36
QY 170 AGGTGCTCCAGATTGTCAACAGGACCGGATGACTTGAAGCAAAAGGATGCGAGTCT 229
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

Db 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACC-----AAGCTCAGTATGATCTCTCGTTGT-----GTCTATGAT 271
 Db 57 ArgCysAsnLeuLeuLysValGluGluGluGluCysGluGluGluGluGluProArg 76
 QY 272 CTTCTGA---GGACACACTGGCACCAACCAACAGCTTCCCTCCAGGGGAGCGGACACGT 328
 Db 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyGluLysGluGluAspGlu 96
 QY 329 GCGCGCAACCCGAGACTACGATGATGACCGC---CGTCAACCCCGAGAGAGGNA--- 382
 Db 97 GlyGluGlnProArgProPheProPheProArgProArgGlnProArgGlnProArgGluGlu 116
 QY 383 -----CGAGGCGCATCG 394
 Db 117 HisGluGlnLysGluGluHisGluTyrHisArgLysGluGluLysHisGlyGlyLysGly 136
 QY 395 GGA-----CCAGCTGGAGCCGAGGAGCGGTGAA 421
 Db 137 SerGluGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156
 QY 422 AGAAGA---GACTGGAGACACCAAGAGAGATGGAGCGGACCAAGTATCATCAGCAG 478
 Db 157 GluGluLysHisGluTyrGlnHisLysGlnGluLysHisGlnGlyLysGluSerGluGlu 176
 QY 479 CCACGGAAATAAGGCCCGAAGGAAGAGAGAACAAAGAG---TGGGGAACACCCAGGT 535
 Db 177 GluGluGluAspGlnAspGluAspGluGluGlnAspLysGluSerGlnLysGluGly 196
 QY 536 AGCCATGTGAGGGAAGAAACATCTCCGAAC-----AACCTTTCTACTTCCCGTCA 586
 Db 197 SerGluSerGlnArgGluProArgArgHisLysAsnLysAsnProPheHisPheAsnSer 216
 QY 587 AGCGGTTAGCACCCGCTACGGGAACCAACAAAGGTAGGATCGGGTCTCGCAGAGGTTT 646
 Db 217 LysArgPheGlnThrLeuPheLysAsnGlnTyrGlyHisValArgValLeuGlnArgPhe 236
 QY 647 GACCAAGGTCAAGGCAGTTTCAGAACTCCAGAACTCCAGATTCAGTATGTGCAGATCAGGCG 706
 Db 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgLleuGluPheAsnSer 256
 QY 707 AACCTTAACACTCTTCTTCCCAAGCACGCTGATGCTGATAACATCTCTGTTATCCAG 766
 Db 257 LysProAsnThrLeuLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValleLeu 276
 QY 767 CAAGGCAACCCACCTGACCGTAGCAATGGCAATAACAGAAAGAGCTTTATCTTGAC 826
 Db 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuGln 296
 QY 827 GAGGGCCATGCACTCAGAATCCCATCCGGTTTCATTTCTCTACATCTTGAACCGCCATGAC 886
 Db 297 SerGlyAspAlaLeuArgValProAlaGlyThrTyrTyrValValAsnProAspAsn 316
 QY 887 AACCAAGACCTCAGAGTAGCTAAATCTCCATGCCGTTTAACACACCCCGCCAGTTTGAG 946
 Db 317 AspGluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGlu 336
 QY 947 GATTTCTTCCGGGAGCAGCCGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAAT 1006
 Db 337 SerPhePheLeuSerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerLysAsn 356
 QY 1007 ACCTTGGAGCGCCTTCAATCGGAATTCATAGATACGAGGCGTGTGTAGAGAG 1066
 Db 357 IleLeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArg 376
 QY 1067 AATGCAGGAGGTGAGCAA---GAGGAGAGAGCGGCAGAGCGATGCGAGTAGT 1123
 Db 377 GluGluGlyGlnGlnGlnGlyGluGluArgLeuGln----- 388
 QY 1124 GAGAAACAATGAGGAGTATGATCAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAG 1183
 Db 389 -----GluSerValIleValGluIleSerLysLysGlnIleArgGluLeuSerLys 405

QY 1184 CACGCTAAATCGTCTCAAAGAAAGGCTCCGAAGAGAGGGAGATATCACCAACCCAATC 1243
 Db 406 HisAlaLysSerSerSerArgLysThrIleSerSerGlu-----AspLysProPhe 422
 QY 1244 AACTTGAGAGAAGGCGAGCCGATCTTCTTAACAACTTTGGGAAGTTATTTAGGTGAAG 1303
 Db 423 AsnLeuArgSerArgAspProIleTyrSerAsnLysLeuGlyLysLeuPheGluIleThr 442
 QY 1304 CCAGACAAGAAAGACCCCGAGCTTCAGGACCTCGGACATGCTCACCTGTCTAGAGATC 1363
 Db 443 ProGlu---LysAsnProGlnLeuArgAspValPheLeuSerValValAspMet 461
 QY 1364 AAAGAGGAGCTTTGATGCTCTCCACACTTCAACTCAAGGCCATGTTATCTCTCGTC 1423
 Db 462 AsnGluGlyAlaLeuPheLeuProHisPheAsnSerLysAlaIleValValLeuValle 481
 QY 1424 ACAAAGGAACCTGGAACCTTGAACTCTGGTGTGTAAGAAAGAGACAACAGAGGGA 1483
 Db 482 AsnGluGlyGluAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnArgGln 500
 QY 1484 CGCGCGGAAGAGAGGAGGAGGACGAGACGAGAGGAGGGAAGCTAACAGAGAGGTGCGT 1543
 Db 501 GlnGlnGluGluGlnPro-----LeuGluValArg 510
 QY 1544 AGGTACACAGCGAGGTTGAAGGAAGCGGATGTTTCATCATGCCAGCAGCTCATCCAGTA 1603
 Db 511 LysTyrArgAlaGluLeuSerGluGlnAspIlePheValIleProAlaGlyTyrProVal 530
 QY 1604 GCCATCAACGCTTCTCCGAACCTCATCTGTTGGTTCGTTATCAACGCTGAAACAAAC 1663
 Db 531 ValValAsnAlaThrSerAspLysAsnPhePheAlaPheGlyIleAsnAlaGluAsnAsn 550
 QY 1664 CACAGAATCTTCTTGCAGGTGATAAGCAATGCTGATAGACCAGTAGAGAGAGCAAGCG 1723
 Db 551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570
 QY 1724 AAGGATTTAGCATTCCTTGGGTGGGTGAAAGTTGAGAAGCTCATCAAAACCCAGAG 1783
 Db 571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590
 QY 1784 GAATCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCCTCTGAG 1843
 Db 591 GluSerTyrPheValAspAlaGlnProGln----- 600
 QY 1844 AAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGGAAGGTCCTCCTC 1903
 Db 601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613
 QY 1904 CTTTCAATTTTGAAGGCTTTT 1924
 Db 614 SerSerIleLeuArgAlaPhe 620
 PRELIMINARY; PRT; 533 AA.
 Q6EBC1
 AC Q6EBC1; (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Beta-conglutinin.
 OS Lupinus albus (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Monteiro S.A., Freitas R.M., Teixeira A.N., Ferreira R.B.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY500372; AAS97865.1; --
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.

RESULT 8

Q6EBC1

ID Q6EBC1

AC Q6EBC1;

DT 25-OCT-2004

DT 25-OCT-2004

DT 25-OCT-2004

DE Beta-conglutinin.

OS Lupinus albus

OC Eukaryota;

OC Spermatophyta;

OC eurosids I;

OX NCBI_TaxID=3870;

RN [1]

RP SEQUENCE FROM N.A.

RA Monteiro S.A.,

RL Submitted (DEC-2003)

DR EMBL; AY500372;

DR GO; GO:0045735;

DR InterPro; IPR006045;

Db	219	AspAsnGlnLysLeuArgValValLysLeuAlaIleProIleAsnAsnProGlyTyrPhe	238
QY	944	GAGGATTTCTCCGGCGACGACCGAGACCAATCATCTTACCTGCAGGCTTCAGCAGG	1003
Db	239	TyrAspPheTyrProSerSerThrLysAspGlnGlnSerTyrPheSerGlyPheSerArg	258
QY	1004	AATACGTTGGAGCCGCTTCAATGCGGAATTCATGAGATACGAGGGTGCTGTTA	1060
Db	259	AsnThrLeuGluAlaThrPheAsnThrArgTyrGluLuiIleGlnArgIleLeuGly	278
QY	1061	---GAGAGANAATCCAGGAGTGAGCAAGAGGAGGCGGAGCGCATCGAGTACTCGG	1117
Db	279	AsnGluAspGluGlnGluTyrGluGluGlnArgArgGlyGlnGluGln	294
QY	1118	AGTAGTGAGAAACAATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGACGACGTTGAAGAACTT	1177
Db	295	---SerAspGlnAspGluGlyValIleValIleValSerLysGlnIleGlnLysLeu	313
QY	1178	ACTAAGCAGCTAAATCCGTCTCAAAGAAAGGCTCCGAAGAAGAGGAGATATACCAAC	1233
Db	314	ThrLysHieAlaGlnSerSerSer-----GlyLysAspLysProSerAsp---SerGly	330
QY	1238	CCAATCAACTTGAGAGNAGCGGAGCCGATCTTCTTAAACAACTTTCGGAAGTATTGTGAG	1297
Db	331	ProPheAsnLeuArgSerAsnGluProIleTyrSerAsnLysTyrGlyAsnPheTyrGlu	350
QY	1298	GTGAAGCCAGACAAGAAGAACCCACAGCTTCAGAGCTTGACATGATGCTCACTGTGTA	1357
Db	351	IleThrProAsp---ArgAsnProGlnValGlnAspLeuAsnIleSerLeuThrTyrIle	369
QY	1358	GAGATCAAGAAGAGGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTC	1417
Db	370	LysIleAsnGluGlyAlaLeuLeuLeuProHisTyrAsnSerLysAlaIleTyrValVal	389
QY	1418	GTGCTCAACAAGGAAGTGGAACTTGAACCTTGCTGCTGTGTAGAAAGAGCAACAACAG	1477
Db	390	ValValAspGluGlyGluGlyAsnTyrGluLeuValGlyIleArgAspGlnGlnArgGln	409
QY	1478	AGGGGACGGCGGAAGAAAGAGGAGGACGAAGACGAAGAGAGAGGAACTAACAGAGAG	1537
Db	410	-----GlnAspGluGlnGluLysGlu-----GluGlu	419
QY	1538	GTCCGTAGGTACACAGCGAGTTGAAGGAGCGCATGTGTTTCATATGCCAGAGCTCAT	1597
Db	420	ValIleArgTyrSerAlaArgLeuSerGluGlyAspIlePheValIleProAlaGlyTyr	439
QY	1598	CCAGTAGGCATCAACGCTTCCCGAACCTCCATCTCTCTGGCTTCGGTATCAACGCTGAA	1657
Db	440	ProIleSerIleAsnAlaSerSerAsnLeuArgLeuLeuGlyPheGlyIleAsnAlaAsp	459
QY	1658	AACAACACAGAAATCTTCTTGAGGTGATAAGGACAAATGTGTATAGACCATAGAGAAG	1717
Db	460	GluAsnGlnArgAsnPheLeuAlaGlySerLysAspAsnValIleArgGlnLeuAspArg	479
QY	1718	CAAGCGAAGATTTAGCATTCCTCGGTCCGGTGAAACAAGTTGAGAGCTCATCAAAAC	1777
Db	480	AlaValAsnGluLeuThrPheProGlySerAlaGluAspIleGluArgLeuIleLysAsn	499
QY	1778	CAGAAGGAATCTCATTGAGTGCTGCTCTCATCTCAATCTCAATCTCAATCTCGTCT	1837
Db	500	GlnGlnGlnSerTyrPheAlaAsnGlyGlnProGlnGlnGlnGlnGln	515
QY	1838	CCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAAACCAAGGAGGAAAGG	1897
Db	516	-----GlnGlnSerGluLysGluGlyArgArgGly	525
RESULT 9			
Q948Y0			
ID	Q948Y0	PRELIMINARY;	PRT; 621 AA.
AC	Q948Y0;		
DT	01-DEC-2001	(TrEMBLrel. 19, Created)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	


```

Db      543 PheGlyIleAsnAlaGluAsnAnGlnArgAsnPheLeuAlaGlySerIysAsnVal 562
QY      1700 ATGACACAGATAGAGAGCAAGCAAGATTTAGCATTCCTCGGTGGGTGAACAAGTT 1759
Db      563 IleserGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaIysAspIle 582
QY      1760 GACAAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAATCTCAA 1819
Db      583 GluAsnLeuIleIysSerGlnSerGluSerTyrPheValAspAlaGlnProGln----- 600
QY      1820 TCTCAATCTCCGTCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGATCAAGAGGAGGAA 1879
Db      601 -----GlnIysGluGluGly 605
QY      1880 AACCAAGAGGAGGAGGAGGTCACCTCTTCAATTTTGAAGCTTTT 1924
Db      606 AsnLysGlyArgIysGlyProLeuSerSerIleLeuArgAlaIaphe 620

RESULT 10
Q94LX2 PRELIMINARY; PRT; 605 AA.
AC Q94LX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-conglycinin alpha subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21327318; PubMed=11434464;
RA Yoshino M., Kanazawa A., Tsutsumi K., Nakamura I., Shimamoto Y.;
RT "Structure and characterization of the gene encoding alpha subunit of
RT soybean beta-conglycinin.";
RL Genes Genet. Syst. 76:199-105(2001).
DR EMBL; AB051865; BAB56161.1; -.
DR PIR; S20007; S20007.
DR HSSP; P25974; 1IPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin_region.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RMLC_like_cupin.
DR Pfam; PF00190; Cupin; 2.
SQ SEQUENCE 605 AA; 70306 MW; 8ACB6F8532662984 CRC64;

Alignment Scores:
Pred. No.: 1,04e-75 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.37% Conservative: 108
Best Local Similarity: 43.05% Mismatches: 173
Query Match: 35.99% Indels: 96
DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q94LX2 (1-605)
QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTGTCTCGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuGluGlyValPheLeuAlaSerVal 19
QY 110 TTGCAACGATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGCGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProIysHisAsn 36
QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGATGACTTGACCAAGCAAGGCATCGCGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCCAGTATGAT-----CCTCGTTGT 262

```

```

Db      57 ArgCysAsnLeuLeuIysValGluLysGluGluCysGluGluGlyGluIleProArg--- 75
QY      263 GTCTATGATCTCTCGAGGACACACTGGCACCAACCAACAGTTTCCCTCCAGGGGAGCG 322
Db      76 -----ProArgProArg-----ProGlnHisProGlu 84
QY      323 ACACGTGGCGCCCAACCCGAGACTACGATGATGAC----- 358
Db      85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluGlnProArgProIle 104
QY      359 -----CGCGCTCAACCCCAAGAGAG-----GAAGAGGAGCCGA 391
Db      105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY      392 TGG-----GGACAGCTGACCCGAGGAGCGTGAAGAGAGAA 430
Db      125 TrpProArgLysGluLysArgGlyGluLysGlySerGluGluAspGluAspGlu 144
QY      431 GACTGGAGACCAACCAAGAGAGAT-----TGAGGCGACCAAGTCAATCAGCAGCCACGG 484
Db      145 AspGluGlnAspGluArgGlnPheProPheProArgProPheHisGlnLysGluGlu 164
QY      485 AAATAAGAGCCCAAGAGAGAGAGAAAGAGTGGGGAACACACAGGTAGCCATGTG 544
Db      165 ArgLysGlnGluGluAspGluAspGluGlnGlnArgGluSerGluSerGluAsp 184
QY      545 AGGGAA-----GAAACATCTCGGAACCAACCTTTCTCTCCGTCAAGCGGTTTAGC 598
Db      185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY      599 ACCCGCTACGGAACCAACCAAGTAGGATCCGGTCTCGAGAGTTTGACCAAGGTCA 658
Db      205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY      659 AGCGAGTTTCAATCTCCAGATCTCCAGATCACCGTATGTGCAGATCGAGGCCCAAACTAACCT 718
Db      225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY      719 CTGTGTTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTCAGCAAGGCGGCAAGCC 778
Db      245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY      779 ACCGTGACCGTAGCAATGGCAATTAACAAAGAGCTTTAATCTTCAGCAGGCGCCATGCA 838
Db      265 IleLeuSerLeuValAsnAsnAspAspAspSerTyrArgLeuGlnSerGlyAspAla 284
QY      839 CTCAGATCCCATCCCGTTTTCATTTCTACATCTTGAACCGCATGACAAACACAGAACCTC 898
Db      285 LeuArgValProSerGlyThrThrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY      899 AGAGTAGCTAAATCTCCATGCCCTTAACACACCCGCGCAGTTTGTAGGATTTCTCCCG 958
Db      305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY      959 GCGACGCGGAGACCAATCATCTTCTGAGGCGTTCAGCGAATACGTTGGAGGCC 1018
Db      325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY      1019 GCCTCAATTCGCAATTCGAATGAGATACGAGGGTGTCTTGAAGAGAATCGAGAGGT 1078
Db      345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY      1079 GAGCAAGAGGAGAGGAGGAGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGA 1138
Db      364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
QY      1139 GTGATAGTCAAGTGTCAAGGAGCACGCTTGAAGAACTTACTAAGCAGCTAAATCCGTC 1198
Db      375 ValIleValGluIleSerLysGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY      1199 TCAAGAAAGGCTCCGGAAGAGGAGATATCACCAACCCCAATCAACTTGAGAGAAAGC 1258
Db      395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411

```

```
QY 1259 GAGCCCGATCTTTCTTAACAACCTTTGGGAAGTTATTTGAGTGAGCCAGACAGGAAGAAC 1318
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 412 AppProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCCAGCTTCAGACCTGGACATGATGCTCACTGCTGTGTGATGATCAAAAGAGAGCTTTG 1378
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCACACTTCAACTCAAGGCGATGTTATGCTGCTGCTCAACAAAGAACTGGA 1438
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 451 LeuLeuProHiePheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTGCTGTAGAAAGAGCAACACAGAGGCGCGCGGAGGAAGAG 1498
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAGAGGAGGAAGTAACAGAGAGGTGCTAGGTACACAGCGAGG 1558
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAGCGATGTTTCATCATGCGCAGAGCTCATCCAGTACCCATCAACGCTTCC 1618
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY 1619 TCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACACACAGAAATCTTCCTT 1678
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGACAAATGTATGATACCAAGATAGAGAAGCAAGCGAAGATTTCATTC 1738
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGlnLeuAlaPhe 559
QY 1739 CTGGTGGCTGGTCAACAGTGTAGAGCTCATCAAAACACAGAGAAATTCATCTTGTG 1798
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGTCTGCTCCTCAATCTCAATCTCCTGCTGCTCCTGAGAAAGAGTCTCTGAG 1858
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCCATCTCTTCAATTTGAAG 1918
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 603 AlaPhe 604

RESULT 11
GLCA_SOYBN
ID _GLCA_SOYBN STANDARD; PRT; 605 AA.
AC P13916;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Beta-conglycinin, alpha chain precursor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=91355860; PubMed=2103438;
RA Sebastiani F.L., Farrell L.B., Schuler M.A., Beachy R.N.;
RT "Complete sequence of a cDNA of alpha subunit of soybean beta-
conglycinin."
RL Plant Mol. Biol. 15:197-201(1990).
CC -!- FUNCTION: Seed storage protein. Accumulates during seed
CC development and is hydrolyzed after germination to provide a
CC carbon and nitrogen source for the developing seedling.
```

```
CC -!- SUBUNIT: The alpha-, alpha-, and beta-subunits associate in
CC various combinations to form trimeric proteins.
CC -!- SUBCELLULAR LOCATION: Embryo axis, and cotyledonary membrane-bound
CC vacuolar protein bodies.
CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X17698; CAA35691.1; -.
CC PIR; S14681; FWSYBA.
CC HSP; P25974; IIFU.
CC InterPro; IPR006045; Cupin.
CC InterPro; IPR007113; Cupin region.
CC InterPro; IPR011031; Multihaem cyt.
CC InterPro; IPR011051; RmlC_like_cupin.
CC Pfam; PF00190; Cupin; 2.
CC Glycoprotein; Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 22
FT PROPEP 23 62
FT CHAIN 63 605 Beta-conglycinin, alpha chain.
FT CARBOHYD 261 261 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 517 517 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 605 AA; 70293 MW; CBEBA30506BBEC57 CRC64;

Alignment Scores:
Pred. No.: 1,21e-75 Length: 605
Score: 1281.00 Matches: 285
Percent Similarity: 59.21% Conservative: 107
Best Local Similarity: 43.05% Mismatches: 174
Query Match: 35.96% Indels: 96
DB: 1 Gaps: 18

US-10-728-323-1 (1-2032) x GLCA_SOYBN (1-605)
QY 50 ATGAGAGGAGGAGTTTCTCCACTGATGCTGTCTAGGATCCTTGTCTGCTTCAGTT 109
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCGATGCGCAAGTCATCCTTACCAGAAAGAAACACAGAGAACCCCTGCGCCCG 169
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGGTGCTCCACAGAGTTGTCAACAGAACCGGATGATCTTGAAGCAAAAGCATGCGAGTCT 229
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 57 ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyGluIleProArg--- 75
QY 263 GTCTATGATCTCTCGAGACACATCGGCACCAACCAACCAACCTTCCCTCCAGGGGCGG 322
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACAGTGGCCGCCCAACCCGAGACTACGATGATGAC----- 358
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 85 ArgGluProGlnGlnProGlyGluLysGluAspGluAspGluGlnProArgProIle 104
QY 359 -----CGCGTCAACCCGCAAGACAG-----GAAGGAGCGCA 391
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCACTGGACCGGAGCGGCGTGAAGAGAA 430
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DQ 125 TrpProArgLysGluGluLysArgGlyLysGlySerGluGluGluAspGlu 144
QY 431 CACTGAGACAAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCAATCAGCAGCCACGG 484
```


Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 170

RESULT 13

Q948X9 PRELIMINARY; PRT; 623 AA.
 AC Q948X9; (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 24, Last annotation update)
 DE Beta-conglycinin alpha-subunit.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_taxid=3847;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Fukazawa C.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030839; BAB64304.1; -.

DR HSSP; P25974; IIPJ.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR007113; Cupin region.

DR InterPro; IPR011051; RmlC_Like_cupin.

DR Pfam; PF00190; Cupin; 2.

SQ - SEQUENCE 623 AA; 72475 MW; EBF033DA1316AB9A CRC64;

Alignment Scores:

Pred. No.: 1,28e-74 Length: 623
 Score: 1265.50 Matches: 283
 Percent Similarity: 58.22% Conservative: 110
 Best Local Similarity: 41.93% Mismatches: 179
 Query Match: 35.53% Indels: 103
 DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q948X9 (1-623)

QY 50 ATGAGAGGAGGTTTCTCCATGATGCTGTGCTAGGATCCTTGTCTGCTGCTCAGTT 109
 Db 1 MetArgAlaAspPhePro-----LeuLeuLeuGlyValValPheLeuAlaSerVal 18
 QY 110 TCTGCAACGATGCCAAGTCATCCTTACCAGAAACAGACAGACCCCTCGGCCAG 169
 Db 19 SerValSerPheGlyIleAla-----TyrTrpGluLys---GlnAsnProLysHisAsn 35
 QY 170 AGTGCTCTCAGAGTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCATGCCAGTCT 229
 Db 36 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 55
 QY 230 CGCTGCACC-----AAGTCGAGTATGATCTCTCGTTGTCTATGATCCTCGAGACAC 283
 Db 56 ArgCysAsnLeuLeuLysValGluLysGluGluCys-----GluGluGlyGlu 72
 QY 284 ACTGGACCAACCAACCAAGTTCCCTCCAGGGGAGGACAGCTGGCGCCCAACCCGGA 343
 Db 73 IleProArgProArgProArgProGlnHisProGluArg---GluProGlnGlnProGly 91
 QY 344 GACTACGATGATGACCGC-----CGTCAA 367
 Db 92 GluLysGluLysAspGluAspGluGlnProArgProIleProPheProArgProArgGln 111
 QY 368 CCCGAAGAGAGGAA----- 382
 Db 112 ProArgGlnGluGluGluHisGluGlnArgGluGluGlnGluTrpProArgLysGluGlu 131
 QY 383 ---GGAGGCGATGGGA-----CCAGCT 403
 Db 132 LysArgGlyGlyLysSerGluGluGlnAspGlyArgGluHisProArgProHis 151
 QY 404 GGACCGGAGGCGTCAAGAGAGAAAGACTGGAGACAACCAAGAAAGATTGGAGCGGA 463
 Db 152 GlnProHisAspGluAspGluGluGlnAspGluArgGln-----PheProPheProArg 169

QY 464 CCAAGTCATCAGCAGCCAGCGAAATTAAG-----CCCGAA 499
 Db 170 ProProHisGlnLysGluSerGluArgLysGlnGluAspGluGluGln 189
 QY 500 GGAAGAGAGGAGAAACAGAGTGGGAAACACACAGGTAGCATGTGAGGGAAGAACATCT 559
 Db 190 GlnArgGluSerGluGluSerGluSerGluArgGlnArgGluLeuArgHisLys 209
 QY 560 CGAAACAACCTTTCTACTCCCGTCAAGCGGTTTAGCCCGCTACCGGAAACCAAAAC 619
 Db 210 AsnLysAsnProPheHisGlySerAsnArgPheGluThrLeuPheLysAsnGlnTyr 229
 QY 620 GGTAGATCCGGTCTCGCAGAGTTTACCAAGGTCAAGGCAGTTTTCAGATCTCCAG 679
 Db 230 GlyArgIleArgValLeuGlnArgPheAsnGlnArgSerProGlnLeuAsnLeuArg 249
 QY 680 AATCACCGTATTGTGCAGATCGAGCCAAACCTAACACTCTTGTCTTCTCCCAAGCAGCT 739
 Db 250 AspTyrArgIleLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProAsnHisAla 269
 QY 740 GATGCTGATAACATCCTTGTATTATCCAGCAAGGGCAGCCCGTACCGTACCGAATGCG 799
 Db 270 AspAlaAspTyrLeuIleAlaIleLeuAsnGlyThrAlaIleLeuSerLeuValAsnAsn 289
 QY 800 AATAACAGAAAGCTTTAATCTTCAGAGGCCCATGCATCAGATCCCATCCCGTTTC 859
 Db 290 AspAspArgAspSerTyrArgLeuGlnSerGlyAspAlaLeuArgValProSerGlyThr 309
 QY 860 ATTTCTCATCTTGAACCGCCATGACCAACCAACCTCAGAGTAGTAAATCTCCATG 919
 Db 310 ThrTyrTyrValValAsnProAsnAsnGluAsnLeuArgLeuIleThrLeuAlaIle 329
 QY 920 CCCGTTAACACACCCGCGCAGTTTTCAGGATTTCTCCCGCGCAGCAGCGACCAATCA 979
 Db 330 ProValAsnLysProGlyArgPheGluSerPhePheLeuSerSerThrGluAlaGlnGln 349
 QY 980 TCCTACTTGCAGGCTTTCAGCAGGAATAGTTGGAGGCGCCTTCAATCGGAAATCAAT 1039
 Db 350 SerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAlaSerTyrAspThrLysPheGlu 369
 QY 1040 GAGATCGAGGCTGTGTAGAAAGATGTCAGGAGGTGAGCAAGAGAGAGAGAGAGAG 1099
 Db 370 GluIleAsnLysValLeuPheSerArgGluGly-----GlnGlnGlnGlyGlu 386
 QY 1100 AGCGGATCGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTGTAGTCAAGTCAAG 1159
 Db 387 GlnArgLeu-----GlnGluSerValIleValGluIleSerLys 399
 QY 1160 GAGCAGTGAAGAACTTACTAAGCAGCCTAATCGTCTCAAGAAAGGCTCCGAAAGAA 1219
 Db 400 GluGlnIleArgAlaLeuSerLysArgAlaLysSerSerSerArgLysThrIleSerSer 419
 QY 1220 GAGGAGATATCACCACCACTCAACTTGAGAAAGGCGAGCCGATCTTCTTAACAAC 1279
 Db 420 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 436
 QY 1280 TTTGGGAAGTTATTTGAGGTGAAGCAGACAAAGAACCCCGAGCTTCAGGACCTGGAC 1339
 Db 437 LeuGlyLysPhePheGluIleThrProGlu---LysAsnProGlnLeuArgAspLeuArg 455
 QY 1340 ATGATGTCACCTGTGTAGAGATCAAGAGAGGCTTTGATGCTCCACACTTCACTCA 1399
 Db 456 IlePheLeuSerIleValAspMetAsnGluGlyAlaLeuLeuLeuProHisPheAsnSer 475
 QY 1400 AAGGCGCTGTTATCGTCTCTCAACAAAGAACTGGAAACCTTGAACCTCGTGGCTGTA 1459
 Db 476 LysAlaIleValIleLeuValIleAsnGluGlyAspAlaAsnIleGluLeuValGlyLeu 495
 QY 1460 AGAAAGAGCAACACAGAGGGGACCGCGGGAAGAGAGGAGACGACGAAGAGAGAG 1519
 Db 496 ---LysGluGlnGln-----GluGluGlnGln 504


```

QY 1520 GAGGGAAGTAACAGAGAGGTGGTAGGTACACAGCGAGGTGGAAGGCGATGTGTC 1579
Db 505 GluGluGlnProLeuGluValArgAlaGlySerGluGluLeuSerGluGlnAspIlePhe 524
QY 1580 ATCATGCCAGCAGCTCATCAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTTGGC 1639
Db 525 ValIleProAlaGlyTyrProValValValAsnAlaThrSerAsnLeuAsnPheAla 544
QY 1640 TTCGGTATCAACGGCTGAAACCAACACACAACTTCCTTCGAGGTGATAAGGACAAATGTG 1699
Db 545 IleGlyIleAsnAlaGluAsnGlnArgAsnPheLeuAlaGlySerGlnAspVal 564
QY 1700 ATAGACCATAGAGAACGAGCAAGGATTTAGCATTCCTCGGTGCGGTGAACAAGTT 1759
Db 565 IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaGlnAlaVal 584
QY 1760 GAGAGCTCATCAAAACCCAGAGGAATCTCACTTTGTGAGTCTCGTCTCAATCTCAA 1819
Db 585 GluLysLeuLeuLysAsnGlnArgGluSerTyrPheValAspAlaGlnProLysLys--- 603
QY 1820 TCTCAATCTCCGTCTCCTGAGAAAGAGTCTCCTGAGAAAGAGATCAAGAGGAGGAA 1879
Db 604 -----LysGluGluGly 607
QY 1880 AACCAAGGAGGAAGGTCCATCTCTTCAATTTGAAGGCTTTT 1924
Db 608 AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe 622

RESULT 14
Q41674 PRELIMINARY; PRT; 545 AA.
ID Q41674;
AC Q41674;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conviclin precursor.
OS Vicia marbonensis (Narbonne vetch).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
OX NCBI_TaxID:3912;
RN [1]_TaxID:3912;
RP SEQUENCE FROM N.A.
RC - TISSUE=Cotyledon;
RA Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 271986; CA996513.1; -.
DR HSSP; P25974; 1IPD.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC_Ilike_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 545 AA; 62810 MW; 459A876F92F5A87E CRC64;

Alignment Scores:
Pred. No.: 3,56e-74 Length: 545
Score: 1258.50 Matches: 279
Percent Similarity: 58.18% Conservative: 98
Best Local Similarity: 43.06% Mismatches: 150
Query Match: 35.33% Indels: 121
DB: 2 Gaps: 16

US-10-728-323-1 (1-2032) x Q41674 (1-545)
QY 38 GTAGCAGGACCAATGAGAGGGAGGTTCCTCCACTGATGCTGTTCGATGGGATCCTGTGC 97
Db 1 MetAlaThrAlaMetLysSerArgPhe---ProValLeuLeuLeuGlyIleIle 19
QY 98 CTGGCTTCAGTTTCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAACAGAGAAC 157

```

```

Db 20 LeuAlaSerLeuCysValThrTyrAlaAsn-----TyrAspGluGlyThrGlu--- 35
QY 158 CCCTCGCGCCAGAGGTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAG 217
Db 35 ----- 35
QY 218 GCATCGAGTCTCGCTCGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGA 277
Db 36 -----ProArgVal----- 38
QY 278 GGACACATCGGCACCAACCAACCAAGTTCCTCCAGGGAGCGGACACAGTGGCGGCCAA 337
Db 39 -----ProGlyGlnArgGluArgGlyArgGln 47
QY 338 CCCGAGACTAGCATGATGACCGCGTCAACCCCGAAGAGGAGGCGGATGGGA 397
Db 48 GluGlyGluLysGluGluLysArgHis-----GlyOluTrpArg 60
QY 398 CCA-----GCTGGACCGAGGAGCGTGAAGA----- 424
Db 61 ProSerHisGluLysGluAlaGlnProGlyArgArgGluArgTrpGluThrSerGluGlu 80
QY 425 -----GAAGAAGACTGGAGA-----CAACCAAGAGAAGATTGGAGGCGACA 466
Db 81 GluGluArgValAspGluGluTrpArgGlySerGlnArgHisGluAsp-----Pro 97
QY 467 AGTCATCAGCACCGCAAAATAAGCCCGAAGAGAGGAGGAGCAAGAGTGG--- 523
Db 98 GluGluArgAlaArgGluArgTrpArgAlaGluGluArgGluArgArgGlnTrpGlu 117
QY 524 GGAACACCGAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTCTTACTTCCCG 583
Db 118 GluGluLysGluGlySerSerLysSerGlnGluArgArgAsnProPheLeuPheLys 137
QY 584 TCAGGCGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGG 643
Db 138 SerAsnLysPheLeuThrLeuPheGluAsnGluAsnGlyHisIleArgArgLeuGlnArg 157
QY 644 TTTGACCAAGGTCAAGCAGTTTCAGAAATCTCCAGATCAACCGTATTGTGCAGATCGAG 703
Db 158 PheAspLysArgSerAspLeuPheGluAsnLeuGlnAsnTyrArgValuGluTyrArg 177
QY 704 GCCAAACCTAACACTCTTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATC 763
Db 178 AlalysProHisThrIlePheLeuProGlnHisIleAspAlaAspLeuIleLeuThrVal 197
QY 764 CAGCAAGGCAAGCACCGTGACCGTAGCAATGGAATAACAGAAAGAGCTTTAATCTT 823
Db 198 LeuSerGlyArgAlaIleLeuThrValLeuSerProAsnAspArgAsnSerTyrAsnLeu 217
QY 824 GACGAGGSCCATGCACCTCAGAAATCCATCCGTTTTCATTTCTCTACATCTTGAACCGCAT 883
Db 218 GluArgGlyAspThrIleLysLeuProAlaGlyThrThrSerTyrLeuLeuAsnGlnAsp 237
QY 884 GACAACCAAGACCTCAGAGTAGCTAAATCTCCATGCCCGCTTAACACACCCGCGCAGTTT 943
Db 238 AspGluGluAspLeuArgValValAspLeuSerIleSerValAsnArgProGlyLysVal 257
QY 944 GAGATTTCTCCGCGGAGGAGCGGAGACCAACATCATCTTACTTGCAGGGCTTCAGCAGG 1003
Db 258 GluSerPheGlyLeuSerGlySerLysAsnGln-----TyrLeuArgGlyPheSerLys 275
QY 1004 AATAGTTTGGAGGCGCTTCAATCGGAATTCATAGATACGAGGAGTCTGTGTGAA 1063
Db 276 AsnIleLeuGluAlaSerLeuAsnThrLysThrGluThrIleGluLysValLeuLeuGlu 295
QY 1064 GAGAAATGAGGAGGTGAGCAAGAGGAGAGGCGGAGGCGGATCGAGTACTCGGAGTAGT 1123
Db 296 Glu-----ProGlnGlnSerIleGlyGlnLysArgArgSerGlnArgGlnGlu 311
QY 1124 GAGAACAAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGGACGTTGAAGAACTTACTAAG 1183
Db 1124 GAGAACAAATGAAGGAGTGATAGTCAAAGTGTCAAAGGAGGACGTTGAAGAACTTACTAAG 1183

```

Db 312 ThrAsn-----AlaLeuVallySerArgGluGlnValGluGluLeuLysArg 328
QY 1184 CAGCTAAATCCCTCTCAAGAAGAGCTCCGAAGAAGAGGAGATATACCAACCCCAATC 1243
Db 329 LeuAlaLysSerSerLysLysGlyValSerSerGluPheGlu-----ProPhe 345
QY 1244 AACCTCAGAGAAGCGGAGCCGATCTTTCTAACAACCTTTGGGAAGTTATTGAGGTGAAG 1303
Db 346 AsnLeuArgSerGlnAsnProLysTyrSerAsnLysPheGlyLysLeuPheGluIleThr 365
QY 1304 CCAGACAAGAAGAACCCCAAGCTTCAGGACCTCGACATGATGTCACCTGTGTGAGATGC 1363
Db 366 ProGluLysLysTyrProGlnLeuGlnAspLeuAspIlePheValSerSerValGluIle 385
QY 1364 AAAGAAGGAGCTTTGATGTCCTCCCACTTCAACTCAAGGCCATGTTATCTGTCGTC 1423
Db 386 AsnGluGlyGlyLeuMetLeuProHisTyrAsnSerArgAlaIleValIleLeuLeuVal 405
QY 1424 AACAAAGGAAGCTGAAACCTTGAACTCGTGCTGTAAAGAAAGAGCAACAAAGAGGGGA 1483
Db 406 AsnGluGlyLysGlyAsnLeuGluLeuValGlyLeuLysAsnGluGlnGlnGlnArg 425
QY 1484 CGCGCGGAAGAAGAGGAGGACGAAGACGAAGAAGAGGAGGAAGTAACAGAGAGTGCGT 1543
Db 426 GluArgGluAspGluGln-----GlnValGln 434
QY 1544 AGGTACACAGCGAGGTGAAGGAAGCGATGTTTCATCATGTCAGAGAGCTCATCAGTA 1603
Db 435 ArgTyrGluAlaArgLeuSerProGlyAspValValIleIleProAlaGlyHisProVal 454
QY 1604 GCCATCAACGCTTCCCGAATCCATCTCGCTTGGCTTCGTTATCAACCTGAAACAC 1663
Db 455 AlaValSerAlaSerAsnLeuLeuLeuGlyPheGlyIleAsnAlaGluAsnAsn 474
QY 1664 CACAGATCTTCTTCGAGTGATAGGACAATGTGATACACAGATAGAGGAAGCAAGCG 1723
Db 475 GlnArgAsnPheLeuThrGlySerAspAspAsnValIleSerGlnIleGluAsnProVal 494
QY 1724 AAGATTTAGCATTCCTGGTGGGTGAACAAGTTGAGAGCTCATCAAAACCAAGAAG 1783
Db 495 LysGluLeuThrPheProGlySerAlaGlnGluValAsnArgLeuLeuLysAsnGlnGlu 514
QY 1784 GAATCTCATTGTTGAGTGCTGCTCTCAATCTCAATCTCAATCTCGTCTCTCTGAG 1843
Db 515 HisSerHisPheAlaAsnAlaGluPro----- 523
QY 1844 AAAGAGTCTCTCAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCACTC 1903
Db 524 -----GluGlnLysGlyGluGluSerGlnArgLysArgSerProIle 537
QY 1904 CTTTCAATTTTGAAGCTTTTAAC 1927
Db 538 SerSerIleLeuGlyThrPheAsn 545
RESULT 15
Q9M3X6 PRELIMINARY; PRT; 613 AA.
ID AC Q9M3X6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Convicilin precursor.
GN Name=cvc;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid1; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed cotyledon;
RA Casey R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ276875; CAB82855.1; -.
DR HSP; P25974; IIPJ.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC_Like_cupin.
DR Pfam; PF00190; Cupin; 2.
KW Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 30 613 convicilin.
SQ SEQUENCE 613 AA; 72063 MW; 49AED99F6135DD19 CRC64;
Alignment Scores:
Pred. No.: 1-148-73 Length: 613
Score: 1251.00 Matches: 283
Percent Similarity: 57.61% Conservative: 107
Best Local Similarity: 41.80% Mismatches: 155
Query Match: 35.12% Indels: 132
DB: 2 Gaps: 17
US-10-728-323-1 (1-2032) x Q9M3X6 (1-613)
QY 68 CCACCTGATCTGCTGTAGGATCCTGTCTCTGGCTTCA---GTTTCTGCAACGATGCC 124
Db 10 ProLeuLeuLeuLeuLeuGlyIleIlePheLeuAlaSerValValCysValThrTyrAla 29
QY 125 AAGTCATCATCTTACAGAGAAGAAACAGAG---AACCCCTGCGCCAGAGGTGCTCCAG 181
Db 30 Asn-----TyrAspGluGlySerGluProArgValProAlaGlnArg-----Glu 44
QY 182 AGTTGTCAACAGAACCCGATGACTTGAAGAAAAGGATCGAGTCTCGTGCACCAAG 241
Db 45 ArgGlyArgGlnGluGlyGluLysGluGluLysArgHisGlyGluTrpArgProSer--- 63
QY 242 CTCGATGATGATCTCTGTCTGCTCTATGATCCTCGAGGACACACTGGCACCAACCAA 301
Db 64 -----TyrGlu-----LysGluGlu 68
QY 302 CTTTCCCTTCAGGGAGCGGACAGTGTGCCCCCAACCCGAGAGACTACGATGATGACCGC 361
Db 69 AspGluGluGluGlyGlnArgGluArgGlyArgGlnGluGlyGluLysGluLysArg 88
QY 362 CGTCAACCCGAGAGAGAGAGAGGAGCGATCGGAGCCAGCT----- 403
Db 89 His-----GlyGluTrpArgProSerTyrGluLysGlnGluAsp 101
QY 403 ----- 403
Db 102 GluGluGluLysGlnLysTyrArgTyrGlnArgGluLysGluAspGluGluLysGln 121
QY 404 -----GGACGAGGGAG 415
Db 122 LysTyrGlnTyrGlnArgGluLysLysGluGlnLysGluValGlnProGlyArgGluArg 141
QY 416 CGTGAAGAGAGAGAGACTGGAGAGACCAACAGAGAGATTGGAGG----- 460
Db 142 TrpGluArgGluGluAspGluGlnValAspGluGluTrpArgGlySerGlnArgArg 161
QY 461 -----CGACCAAGTCATCAGCAGCCAGCCAGCAAAATAAAG 493
Db 162 GluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLysArgAspArg 181
QY 494 CCCGAAGGAAGAAGAGAGAACAGAGTGGGAAACCAAGGTAGCCATGTGAGGGAGAA 553
Db 182 ArgHisGlnArgGlyGluGluGlu-----GluArgSerSerGluSer 196
QY 554 ACATCTCGGAACAACCTTTCTACTTCCGTCGTCAGGCGGTTTACGACCCGCTACGGGAC 613
Db 197 GlnGluArgArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeuPheGluAsn 216
QY 614 CAAAACGCTAGGATCCGGTCTCTGACAGAGTTTACCAAGAGGTTCAGAGTTTCAAGAT 673
Db 217 GluAsnGlyHisIleArgLeuLeuGlnArgPheAspLysArgSerAspLeuPheGluAsn 236

```
QY 674 CTCAGAATCACGTTATGTGCAGATCGAGGCCAAACCTAACTCTTGTCTTCCCAAG 733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 LeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePheLeuProGln 256
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 CAGCGTGTGCTGATACATCCCTTGTATCCAGCAAGGGCAAGCCACCGTAGCCGTAGCA 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 HisIleAspAlaAspLeuIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeu 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 794 AATGGCAATAACAGAGAGCTTAACTCTGACGAGGGCCATCCACTCAGATCCCATCC 853
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 SerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLysLeuProAla 296
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 854 GGTTCATTCTCTACATCTTGAACCCCATGACCAACAGACCTCAGAGTAGTAAATC 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 GlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgLeuValAspLeu 316
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 914 TCCATGCCCGTTAAACACCGGCCAGTTTGTAGGATTTCTCCCGCGAGCAGCCGAGAC 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 ValIleProValAsnGlyProGlyLysPheGluAlaPheAspLeuAlaLysAsnLysAsn 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 974 CAATCATCTACTTGCAGGGCTTCAGCAGGAATACGTTGAGCGCGCTTCAATCGGAA 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 Gln-----TyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerTyrAsnThrArg 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1034 TTCAATGAGATACGGAGGCTGCTGTAGAGAGAAATGCAGGAGGTGAGCAAGAGGAGAGA 1093
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TyrGluThrIleGluLysValLeuLeuGluGlu-----GlnGluLys 368
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1094 GGGCAGAGCGATGGAGTACTCGGAGTAGTAGACAATCAAGGAGTAGATGCAAAAGTG 1153
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 AspArgLysArg-----ArgGlnGlnGlyGluGluThrAspAlaIleValLysVal 385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1154 TCAAGGAGCAGCTTCAAGAACTTACTAAGCAGCGTAAATCGTCTCAAAGAAAGGCTCC 1213
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 SerArgGluGlnIleGluGluLeuLysLysLeuAlaLysSerSerLysLysSerLeu 405
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1214 GAAGAGAGGGAGATATCAACCACTTGAAGAGAGGGCGAGCCCGATCTTTCT 1273
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 ProSerGluPheGlu-----ProIleAsnLeuArgSerHisLysProGluTyrSer 422
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1274 AACAACTTTGGAGAGTATTGTAGGTGAAGCAGCAGACAAGAACCCCGAGCTTCAGGAC 1333
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 AsnLysPheGlyLysLeuPheGluIleThrProGluLysLysTyrProGlnLeuGlnAsp 442
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1334 CTGGACATCATCTCACCTGTGTAGAGATCAAAGAGGAGCTTGTATGCTCCACACTTC 1393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 LeuAspLeuPheValSerCysValGluIleAsnGluGlyAlaLeuMetLeuProHisTyr 462
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1394 AACTCAAGGCCATGGTTATCGTCGTCGTCAACAAAGGAACCTTGAACCTCGTG 1453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 463 AsnSerArgAlaIleValValLeuLeuValAsnGluGlyLysGlyAsnLeuGluLeuLeu 482
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1454 GCTGTAAGAAAAGACCAACACAGAGGGGACGGCGGGAAGAGAGGAGCAACAGCA 1513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 483 GlyLeuLysAsnGlnGlnGluArg-----GluAspArg 494
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1514 GAAGAGAGGGGAAGTAACAGAGAGGTGCTAGTACACAGCGAGGTGGAAGGAGCGCAT 1573
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 LysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSerProGlyAsp 512
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1574 GTGTTTCATCATCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTG 1633
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 ValValIleIleProAlaGlyHisProValAlaIleThrAlaSerSerAsnLeuAsnLeu 532
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1634 CTGGCTTCGGTATCAACCGTCAAAAACAACACAGAACTTCTTCGAGGTGATAAGGAC 1693
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 LeuGlyPheGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeuSerGlySerAspAsp 552
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1694 AATGTGATAGACAGACAGAGCAAGCGAAGGATTTAGCATTTCCCTGGGTGGGTGAA 1753
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 553 AsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGlySerValGln 572
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1754 CAAGTTGAGAAGCTCATCAAAAAACAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAA 1813
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 GluIleAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaAsnAlaGluPro--- 591
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1814 TCTCAATCTCAATCTCCGTCGTCCTCTCTGAGAAAGAGTCTCCTGAGAAAGAGATCAAGAG 1873
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 -----GluGlnLysGlu 595
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1874 GAGGAAAACCAAGGAGGAAAGGTCCTCTTCAATTTGAAGGCTTTT 1924
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 596 GlnGlySerGlnGlyLysArgSerProLeuSerSerIleLeuGlyThrPhe 612
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 24, 2005, 10:04:41
Job time : 157.355 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 09:03:06 ; Search time 35.1903 Seconds
(without alignments)
8620.958 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3562
Sequence: 1 aataacatatatttcac.....cgttgtggtgtttctcc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat 23082005 124356 29238/app query.fasta_1.4757
-DB=Issued Patents AA -OFFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1.46 -runat 23082005 124356 29238 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3278	92.0	626	4	US-09-106-872A-4 Sequence 4, Appli
2	1700	47.7	335	4	US-09-106-872A-17 Sequence 17, Appli
3	1279	35.9	605	1	US-07-955-905A-24 Sequence 24, Appli
4	1267	35.6	571	1	US-07-955-905A-25 Sequence 25, Appli
5	1068	30.0	410	1	US-07-955-905A-26 Sequence 26, Appli
6	811	22.8	566	1	US-07-955-905A-2 Sequence 2, Appli
7	811	22.8	566	1	US-07-955-905A-22 Sequence 22, Appli
8	788	22.1	421	1	US-07-955-905A-27 Sequence 27, Appli
9	777	21.8	587	1	US-07-955-905A-23 Sequence 23, Appli
10	591.5	16.6	489	4	US-09-424-283-3 Sequence 3, Appli
11	569	16.0	524	4	US-09-424-283-1 Sequence 1, Appli
12	559.5	15.7	448	4	US-09-323-195A-18 Sequence 18, Appli

13	514.5	14.4	409	4	US-09-424-283-4	Sequence 4, Appli
14	513.5	14.4	523	4	US-09-323-195A-17	Sequence 17, Appli
15	501	14.1	444	4	US-09-424-283-2	Sequence 2, Appli
16	183	5.1	762	4	US-09-252-991A-29423	Sequence 29423, A
17	173	4.9	1564	4	US-10-144-198-2	Sequence 2, Appli
18	173	4.9	1564	4	US-10-144-198-4	Sequence 4, Appli
19	166.5	4.7	720	4	US-09-252-991A-21881	Sequence 21881, A
20	164.5	4.6	1266	4	US-09-252-991A-30851	Sequence 30851, A
21	163	4.6	714	4	US-09-949-016-10296	Sequence 10296, A
22	160	4.5	630	4	US-09-252-991A-26324	Sequence 26324, A
23	160	4.5	801	4	US-09-252-991A-29274	Sequence 29274, A
24	160	4.5	1898	1	US-08-056-200-94	Sequence 94, Appli
25	160	4.5	1898	2	US-08-800-644-94	Sequence 94, Appli
26	160	4.5	1898	4	US-09-538-092-1280	Sequence 1280, Ap
27	159	4.5	1043	4	US-09-252-991A-28885	Sequence 28885, A
28	158.5	4.4	1075	4	US-09-252-991A-18387	Sequence 18387, A
29	157	4.4	1780	1	US-08-769-309A-5	Sequence 5, Appli
30	157	4.4	1780	3	US-08-994-570-5	Sequence 5, Appli
31	157	4.4	1781	4	US-09-961-403-13	Sequence 13, Appli
32	157	4.4	2079	4	US-09-949-016-8301	Sequence 8301, Ap
33	155	4.4	754	4	US-09-976-594-375	Sequence 375, App
34	155	4.4	760	4	US-09-949-016-11129	Sequence 11129, A
35	155	4.4	863	4	US-09-252-991A-26099	Sequence 26099, A
36	154	4.3	650	4	US-09-252-991A-20375	Sequence 20375, A
37	153.5	4.3	1233	4	US-09-688-1888-89	Sequence 89, Appli
38	153.5	4.3	1233	4	US-09-291-417D-89	Sequence 89, Appli
39	152.5	4.3	580	4	US-09-248-796A-14473	Sequence 14473, A
40	152	4.3	1162	2	US-08-728-323A-2	Sequence 2, Appli
41	152	4.3	1162	3	US-09-298-568-2	Sequence 2, Appli
42	152	4.3	1162	4	US-09-410-399-2	Sequence 2, Appli
43	152	4.3	1162	4	US-09-894-273-2	Sequence 2, Appli
44	152	4.3	1427	4	US-09-252-991A-27005	Sequence 27005, A
45	151.5	4.3	917	4	US-09-252-991A-25101	Sequence 25101, A

ALIGNMENTS

RESULT 1

US-09-106-872A-4
; Sequence 4, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannion, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
; FEATURE:
; OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 1
; OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 2
; OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 3
; OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 4
; OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,
; OTHER INFORMATION: peptide 5
; OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,

QY 1550 ACAGCGAGGTGAAGGAGCGATGCTTTCATCATGCGAGCGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGCAATCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATAGGCAATGTGATAGACCAAGAGAGAGAGAGAGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTGGTGGGTGAACAAGTTGAGAGCTCATCAAAACACAGAGGATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTGCTCTCAATCTCAATCTCCGTCGTCTCTCGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAGAAACAAGAGGAGAGGTCCTCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTGAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 2

US-09-106-872A-17

; Sequence 17, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-17

Alignment Scores:

Pred. No.: 8,73e-156 Length: 335
Score: 1700.00 Matches: 335
Percent Similarity: 99.41% Conservative: 0
Best Local Similarity: 99.41% Mismatches: 0
Query Match: 47.73% Indels: 2
DB: 4 Gaps: 1

US-10-728-323-1 (1-2032) x US-09-106-872A-17 (1-335)

QY 917 ATGCCGTTAAACACCCGCCAGTTTGGAGATTCTTCCCGGCGAGCGAGCCAGACCA 976
Db 1 MetProValAsnThrProGlyGlnPheGluAspPheProAlaSerSerArgAspGln 20
QY 977 TCATCTTACTTGCAGGCTTCAGCAGGAATACGTGAGGCGCCCTTCAATGCGGAATTC 1036
Db 21 SerSerTyrlleuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPheAsnAlaGluPhe 40

QY 1037 AATGATACGAGCGGTGCTTGTAGAGAGAAATGCAGGAGGTGAGCAAGAGGAGAGGG 1096
Db 41 AsnGluIleArgArgValLeuLeuGluGluAsnAlaGlyGlyGluGlnGluGluArgGly 60
QY 1097 CAGAGCGGATGAGTAGTACTCGGAGTACTGAGAAACAATCAAGAGGAGTGTAGTCAAAAGTGCTCA 1156
Db 61 GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIleValLysValSer 80
QY 1157 AGGAGCAGCTTGAAGAATCTTAAAGCAGCTTAAATCCGTCTCAAGAAAGAGCTCCGAA 1216
Db 81 LysGluHisValGluGluThrLysHisAlaLysSerValSerLysLysGlySerGlu 100
QY 1217 GAAGGCGGAGATATCACCAACCAATCAACTGAGAGAGGCGAGCCCGATCTTTTCAAC 1276
Db 101 GluGluGlyAspIleThrAsnProIleAsnLeuArgGluGluProAspLeuSerAsn 120
QY 1277 AACTTTGGGAAGTTATTTGAGGTGAAGCCAGACAAAGAAACCCCGAGCTTTCAGGACCTG 1336
Db 121 AsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGlnLeuGlnAspLeu 140
QY 1337 GACATGATCTCATCTGTGTAGAGATCAAGAAAGAGAGCTTTGTATGCTCCACACTTCAAC 1396
Db 141 AspMetMetLeuThrCysValGluIleLysGluGlyAlaLeuMetLeuProHisPheAsn 160
QY 1397 TCAAGGCGCATGGTTATCGTCTGTCACAAAGGAACCTGGAAACCTTGAACCTCGTGGCT 1456
Db 161 SerLysAlaMetValIleValValAsnLysGlyThrGlyAsnLeuGluLeuValAla 180
QY 1457 GTAAGAAAAAGAGCAACAACAGAGGGGACCGCGGAAAGAGAGAGGACCAAGACGAAGAA 1516
Db 181 ValArgLysGluGlnGlnGlnArgGlyArgGluGluGluGluAspGluGlu 200
QY 1517 GAGGAGGGAAGTAAACAGAGAGGTGCTAGGTACAGAGCGAGGTGGAAGAGCGAGTGTG 1576
Db 201 GluGluGlySerAsnArgGluValArgArgTyrThrAlaArgLeuLysGluGlyAspVal 220
QY 1577 TTCATCATGCCAGCGCTCATCCAGTAGCCATCAACGCTTCTCCGCACTCCATCTGCTT 1636
Db 221 PheIleMetProAlaAlaHisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeu 240
QY 1637 GCGTTCGGTATCAACGCTCAAAAACACACAGAAATCTCTCTTCGAGGTGATAGGACAAAT 1696
Db 241 GlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsn 260
QY 1697 GTGATAGACAGATAGAGAAAGCAAGAGGATTTAGCATTCCTCGGTCGGGTGAACAA 1756
Db 261 ValIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGln 280
QY 1757 GTTGAGAGCTCATCAAAACCAAGAGGAATCTCACTTTGTGAGTCTCGTCTCTCAATCT 1816
Db 281 ValGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAla-----GlnSer 298
QY 1817 CAATCTCAATCTCCGTCGTCTCTGAGAAAGAGTCTCTCGAGAAAGAGGATCAAGAGGAG 1876
Db 299 GlnSerGlnSerProSerSerProGluLysGluSerProGluLysGluAspGlnGluGlu 318
QY 1877 GAAAACCAAGAGGAGGAGGTCACCTCTTTCAATTTTGAAGGCTTTTAAAC 1927
Db 319 GluAsnGlnGlyGlyLysGlyProLeuLeuSerIleLeuLysAlaPheAsn 335

RESULT 3

US-07-955-905A-24
; Sequence 24, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS


```

Db      540 AlAGlySerGlnAspAenValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
Qy      1739 CTTGGGTCGGGTGAACAGTTGAGAGCTCATCAAAACAGAGGAATCTCACTTTGTG 1798
Db      560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAenGlnArgGluSerTyrPheVal 579
Qy      1799 AGTGCTCGTCTCAATCTCAATCTCCTCGTCTCTCCTGAGAAAGTCTCTCTGAG 1858
Db      580 AspAlaGlnProLysLys----- 585
Qy      1859 AAAGAGGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
Db      586 -----LysGluGluGlyAenLysGlyArgLysGlyProLysSerSerIleLeuArg 602
Qy      1919 GCTTTT 1924
Db      603 AlaPhe 604

RESULT 4
US-07-955-905A-25
; Sequence 25, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pisum sativum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..571
; OTHER INFORMATION: /note= "Convicilin from P. sativum"
US-07-955-905A-25

Alignment Scores:
Pred. No.: 1..13-113 Length: 571
Score: 1267.00 Matches: 276
Percent Similarity: 56.67% Conservatives: 98
Best Local Similarity: 41.82% Mismatches: 146
Query Match: 35.57% Indels: 140
DB: 1 Gaps: 13

US-10-728-323-1 (1-2032) x US-07-955-905A-25 (1-571)
Qy      68 CCACTGATGCTGTGTAGGATCTTGTCTCGCTTCAGTTTCTGCAAGCATGCCAAG 127
Db      10 ProLeuLeuLeuPheLeuGlyIleIlePheLeuAlaSerValThrTyrAlaAen 29
Qy      128 TCATCCTTACCAGAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
Db      29 ----- 29
Qy      188 CAACAGGAACCGGATGACTTGAAGCAAAAGGATGCGAGTCTCGCTGCACCAAGCTCGAG 247
Db      29 ----- 29

```

```

Qy      248 TATGATCTCGTGTGTCTATGATCTCGAGGACACACTGGCACCACCACCAACCAACGTTCC 307
Db      30 -----TyrAspGluGlySerGluThrArgVal----- 38
Qy      308 CCTCAGGAGGAGCGGACAGCTGCGCCCAACCCGAGAGCTAGCATGATGATCAGCCGCTCAA 367
Db      39 ---ProGlyGlnArgGluArgGlyArgGlnGluGlyGluGlyGluGlyLysArgHis--- 56
Qy      368 CCCCAGAGAGAAAGGAGCGCATGGGACACAGCTGTGGACCCGAGGAGCGT----- 418
Db      57 -----GlyGluTrpArgProSerTyrGluLysGluGluHisGluGlu 70
Qy      418 ----- 418
Db      71 GluLysGlnLysTyrArgTyrGlnArgGluLysLysGluGlnLysGluValGlnProGly 90
Qy      419 -----GAAAGAGAAGAACTGGAGACACCAAGAGAGATTGGAGG----- 460
Db      91 ArgGluArgTrpGluArgGluGluArgGluGlnValGluGluGluTrpArgGlySer 110
Qy      461 -----CGACCAGTCATCAGCAGCCACCG 484
Db      111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLys 130
Qy      485 AAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 544
Db      131 ArgAspArgArgHisGlnArgGluGlyGluGluGlu-----GluArgSer 145
Qy      545 AGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCTCAAGCGGTTTAGCACCCGC 604
Db      146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165
Qy      605 TACGGGAACAAACCGTAGGATCCGGTCTCTGAGAGGTTTGACAAAGGTCAGGCGAG 664
Db      166 PheGluAsnGluAsnGlyHisIleArgArgLeuGlnArgPheAspLysArgSerAspLeu 185
Qy      665 TTTTCAGATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAACAACCTTAACACTCTTGT 724
Db      186 PheGluAsnGluAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205
Qy      725 CTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAAGGCGCAAGCCACCGTG 784
Db      206 LeuProGlnHisIleAspAlaAspLeuIleLeuValValLeuAsnGlyLysAlaIleLeu 225
Qy      785 ACCGTAGCAATGGCAATAACAGAAAGACTTTAATCTTTCAGAGGCGCATGCACCTCAGA 844
Db      226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245
Qy      845 ATCCCATCCGGTTTCATTTCTTCTACATCTTGAACCCGCATGACACACAGCACTCAGAGTA 904
Db      246 IleProAlaGlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
Qy      905 GCTAAATCTCCATGCTCCCGTTAACACACCCGCGCAGTTTGTAGAGATTCTTCCGCGGAGC 964
Db      266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
Qy      965 AGCCGAGACCAATCATCTTTCAGAGGCTTTCAGAGGAATACGTTGAGAGCGCGCTTC 1024
Db      284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerLeu 303
Qy      1025 AATGCGGAATTCATGAGATACGGAGGCTGCTGTAGAGAGATGACAGAGGTGAGCA 1084
Db      304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGlnLysLysProGln 323
Qy      1085 GAGGAGAGAGGCGCAGAGGCGAGTACTCCGAGTAGTGAGAAACAATGAAGAGTGATA 1144
Db      324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
Qy      1145 GTCAAAGTGTCAAAGGAGCACTTCAAGAACTTACTTAAGACGCTAAATCTCTCTCAAG 1204
Db      341 IleLysValSerArgGluGlnIleGluLeuArgLysLeuAlaLysSerSerSerLys 360
Qy      1205 AAGGCTCGAAGAGAGGAGATATCACCAACCCATCACTTGAGAGAGGCGAGCCC 1264

```

```
Db 361 LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
QY 1265 GATCTTTCTAAACACTTTGGGAGTTATTGTAGGTGAAGCCAGACAGAGAACCCCGAG 1324
Db 378 GluTyrSerAsnLysPheGlyLysPheGluIleThrProGluLysLysTyrProGln 397
QY 1325 CTTACGAGCTGACATGATGCTCACCTGCTAGAGATCAAAAGAGAGCTTTGATGCTC 1384
Db 398 LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
QY 1385 CCACACTTCAACTCAAGGCGCATGGTTATCGTCGTCTGCTCAACAAAGAACTCGAAACCTT 1444
Db 418 ProHisTyrAsnSerArgAlaIleValLeuValLeuValAsnGluGlyLysGlyAsnLeu 437
QY 1445 GAATCGTGGCTCTAAGAAAGAGCAACACAGAGGGGAGCGCGGAAGAGGAGGAC 1504
Db 438 GluLeuLeuGlyLeuLysAsnGluGlnGlnGluArg----- 449
QY 1505 GAAGACGAAGAGAGGAGGAAATTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAG 1564
Db 450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
QY 1565 GAAGGCGATGTGTTTCATCATGCGAGCAGCTATCCAGTAGCCATCAACGCTTCTCCGAA 1624
Db 468 ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerSerAsn 487
QY 1625 CTCCTATCTCTGGCTTCGTTATCAAGCTGTAAGAAACCAACACAGATCTTCCTTCAGGT 1684
Db 488 LeuAsnLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
QY 1685 GATAAGCAACATGTGATACACAGATAGAGAGAGCAAGCAAGATTTAGCATTCCTCGG 1744
Db 508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
QY 1745 TCGGGTGAACAAGTTGAGAGTCTATCAAAAACCAAGAAATCTCACTTTGTGAGTGTCT 1804
Db 528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysSerHisPheAlaSerAla 547
QY 1805 CGTCTCAATCTCAATCTCAATCTCGTCTCTCGTCTCTGAGAAAGAGTCTCTGAGAAAG 1864
Db 548 GluPro----- 550
QY 1865 GATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCTCTCTTCAATTTTGAAGGCTTTT 1924
Db 551 GlnLysGluGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 5
US-07-955-905A-26
; Sequence 26, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pisum sativum
```

```
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1-410
; OTHER INFORMATION: /note= "vicilin from P. sativum"
US-07-955-905A-26

Alignment Scores:
Pred. No.: 1,77e-94 Length: 410
Score: 1068.00 Matches: 209
Percent Similarity: 71.64% Conservatve: 84
Best Local Similarity: 51.10% Mismatches: 104
Query Match: 29.98% Indels: 12
DB: 1 Gaps: 5

US-10-728-323-1 (1-2032) x US-07-955-905A-26 (1-410)
QY 542 GTGAGGGAAGAAACATCTCGGAACAAACCTTTTCTACTTCCCGTCGAGCGGGTTTAGCACC 601
Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32
QY 602 CGCTACGGGAACCAAAACGGGTAGGATCCGGTCTCTGACAGAGTTTGACCAAGGTCACAGG 661
Db 33 LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52
QY 662 CAGTTTCAAGATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACTCTT 721
Db 53 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrLeu 72
QY 722 GTTCTTCCCAAGCAGCGTGATGCTGATTAACATCTCTTGTATTATCCAGCAAGGCAAGCCACC 781
Db 73 PheLeuProGlnTyrThrAspAlaAspPheIleLeuValLeuSerGlyLysAlaThr 92
QY 782 GTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTAATCTTTCGAGAGGCCATCGACTC 841
Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112
QY 842 AGAATCCCATCCCGTTTTCATTTCTACATCTTGAACCGCCATGACAAACCAAGCCTCAGA 901
Db 113 LysLeuProAlaGlySerIleAlaTyrPheAlaAsnArgAspAsnGluGluProArg 132
QY 902 GTAGTAAATCTCATGCCCGTTAAACACACCGCGCAGTTTGAGGATTTCTCCCGGGG 961
Db 133 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152
QY 962 AGCAGCCAGACCAATCATCTACTTGCAGGCTTTCAGCAGGAATACGTTGGAGCCCGCC 1021
Db 153 GlyThrGlnAsnGlnLysSerSerLeuSerGlyPheSerLysAsnIleLeuGluAlaIa 172
QY 1022 TTCAATGCGGAATTCAATGAGATACGGAGGTGCTGTTAGAGAGAAATGCGAGGAGGTGAG 1081
Db 173 PheAsnThrAsnTyrGluGluIleGluLysValLeuLeuGluGlnGln-----Glu 189
QY 1082 CAAGGAGAGAGGCGCAGAGCGGATGGAGTACTCGGAGTAGTGAG---AACAAATGAAGA 1138
Db 190 GlnGluProGlnHisArgArgSerLeuLysAspArgGlnGlnIleAsnGluGluAsn 209
QY 1139 GTGATAGTCAAAAGTGTCAAAGGAGCAGCTTCAAGAACTTACTAAGCAGCTAAATCCGTC 1198
Db 210 ValllevalLysValSerArgAspGlnIleGluLeuSerLysAsnAlaLysSerSer 229
QY 1199 TCAAGAAAGGCTCCGAAGAAGAGGAGAGATATACCAACCAACCAATCAACTTGAGGAAGGC 1258
Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246
QY 1259 GAGCCCGATCTTTCTAACAACTTTGGGAGAGTTATTGTAGGTGAAGCCAGACAGAGAAC 1318
Db 247 AsnProIleTyrSerAsnLysPheGlyLysPheGluIleThrProGlu---LysAsn 265
QY 1319 CCCAGCTTCAGGACCTCGACATGATGCTACCTGCTGTGTAGAGATCAAAAGAGGAGCTTTG 1378
Db 266 GlnGlnLeuGlnAspLeuAspIlePheValAsnSerValAspIleLysValGlySerLeu 285
QY 1379 ATGCTCCCACTTCAACTCAAGGCCATGTTATCGTCGTCGTCGTCACAAAGGAAGACTGGA 1438
```

Db 286 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 305
QY 1439 AACCTTTAACTCGTGGCTCTTAAGAAAGAGCAACAACAGAGGCGCGGGAAGAGAG 1498
Db 306 AspPheGluLeuValGlyGlnArgAsnGluAsnGlnGlyLys-----GluAsn 321
QY 1499 GAGGACGAAGACCAAGAGAGGAGGAAGTAACAGAGAGTGCGTAGGTACACAGGAGG 1558
Db 322 AspLysGluGluGlnGluGluGluGluThrSerLysGlnValGlnLeuTyrArgAlaLys 341
QY 1559 TTCAAGGAAGCGCATGTTTCATCATGTCAGACAGCTCATCCAGTAGCCATCAACGCTTCC 1618
Db 342 LeuSerProGlyAspValPheValIleProAlaGlyHisProValAlaIleAsnAlaSer 361
QY 1619 TCCGAATCCATCTGCTTGGCTTCGATCAACGCTGAAACCAACACAGAAATCTTCCTT 1678
Db 362 SerAspLeuAsnLeuIleGlyLeuGlyIleAsnAlaGluAsnAsnGluArgAsnPheLeu 381
QY 1679 GCAGGTGATAAGCAATGTGTATGATAGACCATAGACAGCAAGCAAGCAATTTAGCATTC 1738
Db 382 AlaGlyGluGluAspAsnValIleSerGlnValGluArgProValLysGluLeuAlaPhe 401
QY 1739 CTGGGTCGGGTCAACAAGTTGAGAAG 1765
Db 402 ProGlySerSerHisGluValAspArg 410

RESULT 6
US-07-955-905A-2
; Sequence 2, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-955-905A-2

Alignment Scores:
Pred. No.: 1,75e-69 Length: 566
Score: 811.00 Matches: 195
Percent Similarity: 50.95% Conservative: 127
Best Local Similarity: 30.85% Mismatches: 218
Query Match: 22.77% Indels: 92
DB: 1 Gaps: 19

US-10-728-323-1 (1-2032) x US-07-955-905A-2 (1-566)
QY 65 TCTCCACTGATGCTGTGTAGGATCTTGTCTCGCTTCAGTTCTTCGACGCGATGCC 124
Db 6 SerProPheIleValIleLeuPheSerLeuLeuLeuSerPheAlaLeuLeuCysSerGly 25
QY 125 AGTCTATCACTTACCAAGAAACACAGAACCCCTGCGCCAG-----AGGTGC 175
Db 26 ValSerAlaTyrGlyArgLysGlnTyrGluArgAspProArgGlnGlnTyrGluGlnCys 45
QY 176 CTCGAGATGTCAACAGAA-----CCGATGACTTGAACAAAGCAATGCCAGTCTCG 232
Db 176 CTCGAGATGTCAACAGAA-----CCGATGACTTGAACAAAGCAATGCCAGTCTCG 232

Db 46 GlnArgArgCysGluSerGluAlaThrGluGluArgGluGlnGluGlnCysGluGlnArg 65
QY 233 TGCACCAAGCTCAGTATGATCTCTGTGTGTCTATGATCTCTCGAGGACACACTGCGACC 292
Db 66 CysGluArg---GluTyrLysGluGln----- 73
QY 293 ACCAACCAAGCTTCCCTCCAGGGGACCGGACACGCTGGCGGCGCAACCCCGAGACTACGAT 352
Db 74 -----GlnArgGlnGlnGlnGluGluGluGlnArgGlnTyrGln----- 86
QY 353 GATGACGCGCTCCGACCAAGAGAGAGGAGCGCGCTGGGACCGACTGCGACCGAGG 412
Db 87 -----GlnCysGlnGlyArgCysGlnGluGlnGlnGln-----GlyGlnArg 100
QY 413 GAGCTGAA-----AGAGAAGAACTGGAGACAAACCAAGAGAAAGATTGGAGG----- 460
Db 101 GlnGlnGlnCysGlnArgLysCysTyrGluGlnTyrLysGluGlnGluArgGlyGlu 120
QY 461 CGACCAAGTCATCAGCAGCCCGGAAATAAGCCCGGAGGAGGAGGAGGAGGAGGAGGAG 520
Db 121 HisGluAsnTyrHisAsnHisLysLysAsnArgSerGluGluGluGlnGlnGln--- 139
QY 521 TGGGGAACACACAGGTAGCCATGTCAGGGAAGAAACATCTCGGAACAACCTTCTACTTTC 580
Db 140 -----ArgAsnAsnProTyrTyrPhe 146
QY 581 CCGTCAAGCGG---TTTAGACCCGCTACGGGAACCAACCGGTAGGATCCGGTCTCG 637
Db 147 ProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlnGlyAsnPheLysIleLeu 166
QY 638 CAGAGTTTGACCAAGGTCAGGAGTTTCAGATCTCCAGAAATCCAGAAATCCAGTATTGTGAG 697
Db 167 GlnArgPheAlaGluAsnSerProLeuLysGlyIleAsnAspTyrArgLeuAlaMet 186
QY 698 ATCGAGGCGCAACCTAACACTCTGCTTCTTCCCAAGCAGCTGCTGCTGATTAACATCTT 757
Db 187 PheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAlaIleTyr 206
QY 758 GTTATCCAGCAGGCGCAAGCCAGCTGACCGTAGCAAAATGGCAATAACAAGAAGAGCTTT 817
Db 207 PheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGluSerTyr 226
QY 818 ATCTTTGAGGCGCATGCTCAGATCCCATCGGTTTCATTTCTTCTTACATCTTGAAC 877
Db 227 AsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrValValSer 246
QY 878 CGCCATGACCAACCAACCTCAGAGTAGCTAAATCTCCATCGCTTAAACACACCCGCG 937
Db 247 GlnAspAsnGlnGlyLysLeuThrIleAlaValLeuAlaLeuProValAsnSerProGly 266
QY 938 CAGTTTGAGGATTTCTCCCGCGAGCAGCGAGACCAATCATCTTCTTCTGAGGCGCTTC 997
Db 267 LysTyrGluLeuPhePheProAlaGlyAsnAsnLysProGluSerTyrTyrGlyAlaPhe 286
QY 998 AGCAGGAATACGTTGAGGCGCTTCAATCGGGAATTCATGATGATGATGATGATGATG 1057
Db 287 SerTyrGluValLeuGluThrValPheAsnThrGlnArgGluGlyLysLeuGluLeuLeu 306
QY 1058 TTAGAAGAGATGTCAGGAGGTGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1117
Db 307 -----GluGlnArgGlyGlnLysArg----- 314
QY 1118 AGTAGTGAGAACATGAAGAGGTGATAGTCAAGGTGTCAAAGGAGCAGCTTGAAGAACTT 1177
Db 315 ---GlnGlnGlnGlnGlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIle 333
QY 1178 ACTAAGACGCTAAATCGCTCTCAAGAAAGCGCTCCGAGAAAGAGGAGGAGATATCACAAC 1237
Db 334 SerGlnGlnAlaThrSerProArgHisArgGlyGlyGluArgLeuAla----- 349
QY 1238 CCAATCACTTGAGAGAGGCGCGCTCTTCTTAACAACCTTTTGGGAAGTATTGTAG 1297
Db 350 ---IleAsnLeuLeuSerGlnSerProValTyrSerAsnGlnAsnGlyArgPhePheGlu 368

```
QY 1298 GTGAAGCCACACAGAAGAACCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTA 1357
Db 369 AlaCysProGluAspPheSer---GlnPheGlnAsnMetAspValAlaValSerAlaPhe 387
QY 1358 GAGATCAAAAGAGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTC 1417
Db 388 LysLeuAsnGlnGlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValVal 407
QY 1418 GTCGTCAACAAGGAACTGGAACTTGAACCTGCTGGCTGTAAGAAAAGAGCAACA--- 1474
Db 408 PheValThrAspGlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSer 427
QY 1475 -----CAGAGGGGACGGCGGGAAGAGAGGAGCAGACCAAGCAAGAGAGGGA 1525
Db 428 GlnGlySerGlnSerGlnArgGlnAspArgGlnGlnGluGlnGluSerGluGlnGlu 447
QY 1526 AGTAACAGAGAGGTGGTGTAGGTACACAGCGAGGTGGAAGGAGGCGATGTTCATCATG 1585
Db 448 ThrPheGlyGluPheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAla 467
QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA-----CTCCATCTGCTTGGC 1639
Db 468 ProAlaGlyHisAlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAla 487
QY 1640 TTCGGTATCAACGCTGAAACCAACACAGAACTTCTTTCAGAGGTGATAGGACAAATGTG 1699
Db 488 PheGlyLeuAsnAlaGlnAsnAsnGlnArgIlePheLeuAlaGly---LysLysAsnLeu 506
QY 1700 ATAGACCAGATAGAGACGAGGAAGGATTTAGCATTCCTCGGTGCGGTGAAACAAGTT 1759
Db 507 ValArgGlnMetAspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuVal 526
QY 1760 GAGAAGCTCATCAAAACCAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAATCTCAA 1819
Db 527 AspAsnIlePheAsnAsnProAspGluSerTyrPheMetSerPheSerGlnArgGln 546
QY 1820 TCTCAATCTCCGTCGTCCTCGAAGAGAGTCTCTCGAAGAGAGGATCAAGAGAGGAA 1879
Db 547 -----ArgArgAspGlu 550
QY 1880 AACCAAGAGGAAAGGTCCACTCTTTCAATTTTG 1915
Db 551 ArgArgGly-----AsnProLeuAlaSerIleLeu 560
RESULT 7
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Theobroma cacao
; FEATURE:
; NAME/KEY: Protein
```

```
; LOCATION: 1..566
; OTHER INFORMATION: /note= "67 kD Precursor Protein"
; US-07-955-905A-22
Alignment Scores:
Pred. No.: 1,75e-69 Length: 566
Score: 811.00 Matches: 195
Percent Similarity: 50.95% Conservative: 127
Best Local Similarity: 30.85% Mismatches: 218
Query Match: 22.77% Indels: 92
DB: 1 Gaps: 19
US-10-728-323-1 (1-2032) x US-07-955-905A-22 (1-566)
QY 65 TCTCCACTGATGCTCTGCTAGGATCTTGTCCCTGGCTTCAGTTTCTGCAACGCATGCC 124
Db 6 SerProPheIleValIlePheSerLeuLeuSerPheAlaLeuLeuCysSerGly 25
QY 125 AAGTCATCACCTTACCAAGAAAGAAACAGAACCCCTCGCCCCAG-----AGGTGC 175
Db 26 ValSerAlaTyrGlyArgLysGlnTyrGluArgAspProArgGlnGlnTyrGluGlnCys 45
QY 176 CTCGAGGTGTTCAACAGGAA---CCGATGACTTTGAAGCAAAAGCGATGCGAGTCTCGC 232
Db 46 GlnArgArgCysGluSerGluAlaThrGluGluArgGlnGlnGlnCysGluGlnArg 65
QY 233 TGCACCAAGCTCGAGTATGATCTCGTTGTGTATGATCTCTCGAGGACACACTGGCACC 292
Db 66 CysGluArg---GluTyrLysGluGln-----73
QY 293 ACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGCGCCCAACCCCGAGACTACGAT 352
Db 74 -----GlnArgGlnGlnGluGluLeuGlnArgGlnTyrGln-----86
QY 353 GATGACCCCGCTCAACCCCGAAGAGAGAGAGAGAGGCCCATGGGACACAGCTGGACCGAGG 412
Db 87 -----GlnCysGlnGlyArgCysGlnGlnGlnGln-----GlyGlnArg 100
QY 413 GAGCGTGAA-----AGAGAAGAGACTGGAGACAACCAAGAGAAGATTGGAGG-----460
Db 101 GluGlnGlnGlnCysGlnArgLysCysTyrGluGlnTyrLysGluGlnGluArgGlyGlu 120
QY 461 CGACCAAGTCTACAGACCCACGGAATAAAGCCCGCAAGGAAGAGAGAGAGAAACAAG 520
Db 121 HisGluAsnTyrHisAsnHisLysLysAsnArgSerGluGluGluGlyGlnGln---139
QY 521 TGGGGAACACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTACTTC 580
Db 140 -----ArgAsnAsnProTyrTyrPhe 146
QY 581 CCGTCAAGGCGG---TTTAGCACCGCTACGGGAACCAACCGGTAGGTACCGGTCTCTG 637
Db 147 ProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPheLysIleLeu 166
QY 638 CAGAGGTTTGACCAAGGTCAAGGCAGTTTCAGAAATCTCCAGAATCACCGTATTGTGCAG 697
Db 167 GlnArgPheAlaGluAsnSerProLeuLysGlyIleAsnAspTyrArgLeuAlaMet 186
QY 698 ATCGAGGCCAAACCTAAACACTCTTGTCTTCCCAAGCAGCGTGATGCTGTATAACATCTTC 757
Db 187 PheGluAlaAsnProAsnThrPheIleLeuProHisHisCysAspAlaGluAlaIleTyr 206
QY 758 GTTATCCAGCAAGGCAAGCCAGTACCGTACCGTAGCAATGCAATTAACAGAAAGAGCTTT 817
Db 207 PheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGluSerTyr 226
QY 818 AATCTTGACGAGGCGCATGCTCAGATCCCATCCCGTTCCTTCTACATCTTGTAAC 877
Db 227 AsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrValValSer 246
QY 878 CGCATGACACACCAAGCAACCTCAGAGTAGTAAATCTCCATCCCGCTTAACACACCCGCG 937
Db 247 GlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSerProGly 266
```



```
QY 1067 AATGCAAGGAGTGTGAGCAAGAGGAGAGGAGGAGGAGGAGTGGAGTACTCGGAGTAGTCAG 1126
Db |||:|||||
209 Glu-----GlyGlnGln----- 212
QY 1127 AACATGAAGGAGTGTAGTCAAGTGTCAAGAGGACAGTGTGAAGAACTTACTTAAGCAC 1186
Db |||:|||||
213 -----GluGlyValIleValAsnIleAspSerGluGlnIleProAlaIleLeuValSerIleHis 230
QY 1187 GCTAATCGTCTCAAGAAAGGCTCCGAAGAAGAGGAGGAGATATCACCAACCCATCAAC 1246
Db |||:|||||
231 AlaLysSerSerArgLys-----Ser 238
QY 1247 TTGAGAGAGGAGGAGCCGCTCTTTCTAACTTTTGGAAAGTTATTTGAGGTGAAGCCA 1306
Db |||:|||||
239 LeuSerLysGlnAspAsnThrIleGlyAsnGluPheGlyAsnLeuThrGlu---ArgThr 257
QY 1307 GACAGAAGAACCCCGAGTTCAGGACCTGGACATGATCTCACCTGTGTAGATCAAA 1366
Db |||:|||||
258 Asp-----AenSerLeuAsnValLeuIleSerSerIleGluMetGlu 271
QY 1367 GAAGGAGCTTGATGCTCCACACTCAACTCAAGGCCATGTTATCTGCTGCTCAAC 1426
Db |||:|||||
272 GluGlyAlaLeuPheValProHisTyrTyrSerLysAlaIleValIleLeuValValAsn 291
QY 1427 AAAGGAAGCTCGAAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACGG 1486
Db |||:|||||
292 GluGlyGluAlaHisValGluLeuValGly-----ProLysGlyAsn 305
QY 1487 CGGAAAGAGAGGAGGAGCAAGACGACGAGAGGAGGAGGAAGTAAACAGAGAGTGGTAGG 1546
Db |||:|||||
306 LysGluThrLeuGluTyrGluSer----- 313
QY 1547 TACACAGCAGGTTGAAGAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCC 1606
Db |||:|||||
314 TyrArgAlaGluLeuSerLysAspValPheValIleProAlaIleTyrProValAla 333
QY 1607 ATCAAGCTTCTCCGAACTCCATCTGCTTGGCTTCGATCAAGCTGAAACCAACCAAC 1666
Db |||:|||||
334 IleLysAlaThrSerAsnValAsnPheThrGlyPheGlyIleAsnAlaAsnAsnAsn 353
QY 1667 AGAATCTTCTGCGAGTGAAGGACATGATGATAGACCATGATAGAGAG-----CAA 1720
Db |||:|||||
354 ArgAsnLeuLeuAlaGlyLysThrAspAsnValIleSerSerIleGlyArgAlaLeuAsp 373
QY 1721 GCGAAGGAT-----TTAGCATTCCTGGTGGGTGAACAAGTTGAGAGCTCATC 1771
Db |||:|||||
374 GlyLysAspValLeuGlyLeuThrPheSerGlySerGlyAspGluValMetLysLeuIle 393
QY 1772 AAAAAACGAGAAGGAATCTCACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCCG 1831
Db |||:|||||
394 AsnLysGlnSerGlySerTyrPheValAspAlaHis----- 405
QY 1832 TCGTCTCTGAGAAGAGTCTCTGAGAAGAGGATCAAGAGAGGAGAAACCAAGAGGG 1891
Db |||:|||||
406 -----HisHisGlnGlnGluGlnLysGlyArg 415
QY 1892 AAGGT 1897
Db |||:|||||
416 LysGly 417
```

RESULT 9

```
US-07-955-905A-23
; Sequence 23, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..587
; OTHER INFORMATION: /note= "vicillin from G. hirsutum"
US-07-955-905A-23
```

Alignment Scores:

Pred. No.:	3,53e-66	Length:	587
Score:	777.00	Matches:	190
Percent Similarity:	51.36%	Conservative:	112
Best Local Similarity:	32.31%	Mismatches:	208
Query Match:	21.81%	Indels:	78
DB:	1	Gaps:	15

US-10-728-323-1 (1-2032) x US-07-955-905A-23 (1-587)

```
QY 104 TCAGTTCTGCAACGCATGCCAAGTCATCATCTTACCAGAAAGAAACAGAGAACCCCTGC 163
Db |||:|||||
64 SerCysLysSerGlnTyrGlyGluLysAspGlnGlnArgHisArgProGluAspPro--- 82
QY 164 GCCCAGAGTGCCTCCAGAGTGTCAACAGAA-----CCGATGACTTGAGCA 214
Db |||:|||||
83 ---GlnArgArgTyrGluGluCysGlnGlnGlnCysArgGlnGlnGluArgGlnArg 101
QY 215 AAGGCATCGAGTCTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCT 274
Db |||:|||||
102 ProGlnCysGlnGlnArgCysIleLys----- 110
QY 275 CGAGGACACTGGCACCACCAACCAACGTTTCCCTCCAGGGAGCGGACAGTGGCCGC 334
Db |||:|||||
111 -----ArgPheGluGlnGlnGlnGlnSerGlnArg 121
QY 335 CAACCCGAGACTAGCATGATGAC---CGCGTCACACCCCGAAGAGAGAGAGGAGCCGA 391
Db |||:|||||
122 GlnPheGlnGluCysGlnGlnHisCysHisGlnGlnGlnGlnArgProGluArgLysGln 141
QY 392 TGGGACCCAGCTGGACCCGAGGAGCGTGAAGAGAAAGAGACTGGAGACAAACCAAGAA 451
Db |||:|||||
142 GlnCysValAlaGluCysArgGluArgTyrGlnGluAsnProTyrArgGluArgGlu 161
QY 452 GATTGAGGCGACCAAGTCATCAGCGCCACCGAAATAAGGCCCGCAAGAGAGAGAA 511
Db |||:|||||
162 Glu-----GluAlaGluGlu 167
QY 512 GAACAGAGTGGGAACACACAGGTAGCATCTGAGGGAAGAAAATCTCGGAACACCT 571
Db |||:|||||
168 GluThrGluGluGly-----GluGlnGlnGlnSerHisAsnPro 180
QY 572 TTCTACTTCCCGTCAAGCGGTTTAGCACCGCTACGGGAACCAACCGTAGGATCCGG 631
Db |||:|||||
181 PheHisPheHisArgArgSerPheGlnSerArgPheArgGluGluHisGlyAsnPheArg 200
QY 632 GTCCTGCAAGGTTTGACAAAGGTCAAGCGAGTTTCAGATCTCCAGAAATCACCGTATT 691
Db |||:|||||
201 ValLeuGlnArgPheAlaSerArgHisProIleLeuArgGlyIleAsnGluPheArgLeu 220
QY 692 GTGCAGATCGAGGCCAAACCTTAACACTCTTTCTTCCCAAGACGCTGATGCTGATAAC 751
Db |||:|||||
221 SerIleLeuGluAlaAsnProAsnThrGluValLeuProHisCysAspAlaGluLys 240
```

```
QY 752 ATCTTTGTTATCCAGCAGGCGCAAGCCAGCTGACCGTAGCAAAATGGCAATATACAGAAAG 811
Db 241 ILeTyLeuValThrAsnGlyArgGlyThrLeuThrPheLeuThrHisGluAsnLysGlu 260
QY 812 AGCTTTAATCTTACGAGGCGCATCCAGTCCAGATCCATCCGCTTTCATTCCTCATAC 871
Db 261 SerTyAsnValProGlyValValValArgValProAlaGlySerThrValTyLeu 280
QY 872 TTGAACCGCATGACACCAACCTCAGAGTAGTAAATCTCCATGCCCGTTAAACACA 931
Db 281 AlaAsnGlnAspAsnLysGluLysLeuLeuLeuLeuLeuLeuHisArgProValAsn 300
QY 932 CCGGCGCAGTTTACGAGATTTCTTCCGGCGAGCAGCGAGACCAATCATCTACTTGCAG 991
Db 301 ProArgGlnPheGluGluPheProAlaGlySerGlnArgProGlnSerTyLeuArg 320
QY 992 GGTCTTACAGAGGAATAGTTGGAGGCGCTTCAATGCGGAATTCATGATACGACGAGG 1051
Db 321 AlaPheSerArgGluLeuLeuGluProAlaPheAsnThr-----ArgSer 335
QY 1052 GTCTCTTTAAGAGAAATCAGGAGGTGAGCAAGAGGAGAGAGGCGAGCGATGGAGT 1111
Db 336 GluGlnLeuAspGluLeuPheGlyArgGlnSerHisArgArgGlnGln----- 352
QY 1112 ACTCGGAGTAGTGAGAACAAATGAAGAGTAGTAGTCAAAGTGTCAAAGAGACGTTGAA 1171
Db 353 -----GlyGlnGlyMetPheArgLysAlaSerGlnGlnGlnIleArg 366
QY 1172 GAACCTTACTAAGCAGCTAAATCCGCTCAAGAAAGGCTCCGAAGAGAGGAGATATC 1231
Db 367 AlaLeuSerGlnGluAlaThrSerProArgGluLysSerGlyGluArgPheAla----- 384
QY 1232 ACCAACCCCAATCAACTTGAGAGAGGCGAGCCGATCTTTCTAAACAATTTTGGAGTTA 1291
Db 385 -----PheAsnLeuLeuTyArgThrProArgTy-SerAsnGlnAsnGlyArgPhe 401
QY 1292 TTTGAGGTGAGCCAGACAGAAAGAACCCCGACCTTCAGACCTGACATGATGCTCAC 1351
Db 402 TyrGluAlaCysPro---ArgGluPheArgGlnLeuSerAspIleAsnValThrValSer 420
QY 1352 TGTGTAGATCAAGAGGAGCTTTGATGCTCCACACTTCAACTCAACGAAGCCATGGTT 1411
Db 421 AlaLeuGlnLeuAsnGlnLysSerIlePheValProHisTyAsnSerLysAlaThrPhe 440
QY 1412 ATCGTCTGCTCAACAAAGAACTGGAACCTTGAACCTGCTGCTGTAAAGAAAGACAA 1471
Db 441 ValValLeuValAsnGluGlyAsnGlyTyValGluMetValSerProHisLeuProArg 460
QY 1472 CAACAGAGGGGCGCGGGAAGAGAGGAGGAGCAAGACGAGAGAGGAGGAAGTAAC 1531
Db 461 GlnSerSerPheGluGluGluGluGluGluGlnGlnGlnGluGluGluArg 480
QY 1532 AGA-----GAGGTGGTAGGTACACAGCGAGGTTCAAGGAGCGCATGTTTCATCAG 1585
Db 481 ArgSerGlyGlnTyArgLysIleArgSerGlnLeuSerArgGlyAspIlePheValVal 500
QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCC-----TCCGAACCTCCATCTGCTTGGC 1639
Db 501 ProAlaAsnPheProValThrPheValAlaSerGlnAsnGlnAsnLeuArgMetThrGly 520
QY 1640 TTCGGT-----ATCAACGCTGAAACCAACACAGATCTTCTTCGAGGT 1684
Db 521 PheGlyLeuTyAsnGlnAsnIleAsnProAspHisAsnGlnArgIlePheValAlaGly 540
QY 1685 GATAAGCAATGTATAGACCATAGAGACGAGCGAGGATTTAGCATTCCTCGGG 1744
Db 541 ---LysIleAsnHisValArgGlnTrpAspSerGlnAlaLysGluLeuAlaPheGlyVal 559
QY 1745 TCGGGTGAAACAGTTGAGAGCTCATCAAAACACAG-----AAGGAATCTCACTTTGTAGT 1801
Db 560 SerSerArgLeuValAspGluIlePheAsnAsnAsnProGlnGlnSerTyPheValSer 579
```

```
QY 1802 GCTCGTCTCTCAATCTCAATCTCAA 1825
Db 580 ArgGlnArgGlnArgAlaSerGlu 587
RESULT 10
US-09-424-283-3
; Sequence 3, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424, 283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047, 568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-3
Alignment Scores:
Pred. No.: 2,99e-48 Length: 489
Score: 591.50 Matches: 156
Percent Similarity: 43.90% Conservative: 100
Best Local Similarity: 26.80% Mismatches: 195
Query Match: 16.61% Indels: 131
DB: 4 Gaps: 13
US-10-728-323-1 (1-2032) x US-09-424-283-3 (1-489)
QY 71 CTGATGCTGTTCTAGGGATCCTTGTCTGGCTTCAGTTTCTGCAACGCATGCCAAGTCA 130
Db 16 LeuLeuAlaLeuIleSerAsnLeuAlaLeuGlyLysLeuLysGluThrGluValGlu--- 34
QY 131 TCACCTTACCAGAAAGAACAGAACCCCTCGCCAGAGGTGCTCCAGAGTTGTCAA 190
Db 35 -----GluAspProGluLeuValThrCysLysHisGlnCysGln 47
QY 191 CAGGAACCGAT-----GACTTGAAGCAAAAGGATCGAGTCTCGTGCACCAAGCTC 244
Db 48 GlnGlnArgGlnTyThrGluSerAspLysArgThrCysLeuGlnGlnCysAspSerMet 67
QY 245 GAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGCACCACCAACCGT 304
Db 68 LysGlnGlu----- 70
QY 305 TCCCTCCAGGGGAGCGGACACGTCGCGCCCAACCCGGAGACTACGATGATGACCGCGT 364
Db 70 ----- 70
QY 365 CAACCCCGAAGAGAGAGAGGAGCGCGTGGGACCACTGGGACCGAGGAGCGGTGAAGA 424
Db 71 -----ArgGluLysGlnVal 75
QY 425 GAAGAAGACTGGAGACACCAAGAGAGATTTGGAGGCGGACCAAGTCAATCAGCAGCCACG 484
Db 76 GluGluGluThrArgGluLysGluGlu-----HisGlnGluGln--- 89
QY 485 AAAATAAGCCCGAAGGAAGAGAGAGAAAGAGTGGGGAACACCAAGGTAGCCATGTG 544
Db 89 ----- 89
QY 545 AGGGAAGAAAACATCTCGGAACAACCTTTTACTTTC-----CCGTCAAGGCGGTTTAGCAC 601
Db 90 HisGluGluGluGluAspGluAsnProTyValPheGluGluAspLysAspPheSerThr 109
QY 602 CGCTACGGGAACCAAAACCGTAGGATCCGGGTCTTCAGAGAGGTTTTCACCAAGGTCAAG 661
```

```
Db 110 ArgValGluThrGluGlyGlySerIleArgValLeuLysLysPheThrGluLysSerLys 129
QY 662 CAGTTTCCAGATCTCCAGATCACCCTATGTGTCAGATCAGGCCAAACCTAACACTCTT 721
Db 130 LeuLeuGlnGlyIleGlnAsnPheArgLeuAlaIleLeuGluAlaArgAlaHisThrPhe 149
QY 722 GTTCTTCCCAAGCAGCTGATGCTGATACATCTTGTATTATCCAGCAAGGCGCAAGCCACC 781
Db 150 ValSerProArgHisPheAspSerGluValValLeuPheAsnIleLysGlyArgAlaVal 169
QY 782 GTGACCGTAGCAAAATGCAATAACAGAAAGAGCTTTAATCTTTCAGCAGGCGCCATGCATC 841
Db 170 LeuGlyLeuValArgGluSerGluThrGluLysIleThrLeuGluProGlyAspMetIle 189
QY 842 AGNATCCCATCCGTTTTCATCTTCTACATCTTGAACCGCCATGACNACAGAACCTCAGA 901
Db 190 HisIleProAlaGlyThrProLeuTyrlIleValAsnArgAspGluAsnGlnLysLeuLeu 209
QY 902 GTAGCTAAATCTCCATGCCGCTTAAACACACCCGGCCAGTTTGAGGATTTCTCCCGGGC 961
Db 210 LeuAlaMetLeuHisIleProValSerThrProGlyLysPheGluGluPheGlyPro 229
QY 962 AGCAGCCGAGACCAATCTCTACTTTCAGGGCTTCAGCAGGAATACGTTGAGGCGGCC 1021
Db 230 GlyGlyArgAspProGluSerValLeuSerAlaPheSerTrpAsnValLeuGlnAlaLa 249
QY 1022 TTCATGCGGAATTCATGATGATACGAGGCTGCTGTAGAGAGNATCCAGGAGGTGAG 1081
Db 250 LeuGlnThrProLysGlyLysLeuGluArgLeuPhe 261
QY 1082 CAAGAGGAGAGGCGCAGAGGCGATGCGAGTACTCGGAGTAGTGAGAAACAATGAGGAGTG 1141
Db 262 262
QY 1142 ATAGTCAAGTGTCAAAGAGACGTTGAGAACTTACTAAGCAGCTAAATCCGTTCTCA 1201
Db 269 IlePheLysIleSerArgGluArgValArgAlaLeu 283
QY 1202 AGAAAGGCTCC-----GAAGNAGGCGAGATATCACCAACCCATCAACTTGTAGAGNA 1255
Db 284 LysLysSerSerTrpTrpProPheGlyGlyLysLysLysAlaGlnPheAsnIlePheSer 303
QY 1256 GCGAGCGCCGATCTTTCTAACTTTGGAAGTATTATTTGAGTGAAGCGACAGCAAGAG 1315
Db 304 LysArgProThrPheSerAsnGlyTyrglyArgLeuThrGluValGlyProAspAspGlu 323
QY 1316 AACCCCGAGCTTCAGGACTTGACATGATGCTCACCCTGTGTAGAGATCAAGAAAGAGCT 1375
Db 324 LysSerTrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSer 343
QY 1376 TTGATGCTCCACACTTCACTCAAGGCGCATGGTTATCGTGTCTCAACAAGGAAGCT 1435
Db 344 MetSerThrIleHisTyraSerHisAlaThrLysIleAlaLeuValMetAspGlyArg 363
QY 1436 GGAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACACAGAGGCGCGCGGAAGAA 1495
Db 364 GlyHisLeuGlnIleSerCysProHisMetSerSerArg 376
QY 1496 GAGGAGGACGAAGACGAAGAGGAGGAAGTAACAGAGAGTGCGTTAGGTACACAGCG 1555
Db 377 -----SerAspSerLysHisAspLysSerSerProSerTyraHisArgIleSerAla 393
QY 1556 AGTTTCAAGGAGGCGATGTTTCATCATCGCAGCAGCTCATCCAGTACCCATCAACGCT 1615
Db 394 AspLeuLysProGlyMetValPheValProProGlyHisProPheValThrIleAla 413
QY 1616 TCTCTCGAACTCATCTGCTT-----GGCTTCGGTATCAACGCTGAAACCAACACACAGA 1669
Db 414 SerAsnLysGluAsnLeuLeuIleCysPheGluValAsnValArgAspAsnLysLys 433
QY 1670 ATCTTCTTCGAGGTGATAGGAACAATGTGTATGATAGACCATAGAGNAGCAAGCGAGGAT 1729
```

```
Db 434 PheThrPheAlaGly---LysAspAsnIleValSerSerLeuAspAsnValAlaLysGlu 452
QY 1730 TTAGCATTTCCCTGGGTGCAACAACACTTCAGAGCTCATCAAAACCCAGAAAGAACTCT 1789
Db 453 LeuAlaPheAsnTyraProSerGluMetValAsnGly-----ValSerGluArgLysGluSer 471
QY 1790 CACTTT 1795
Db 472 LeuPhe 473
RESULT 11
US-09-424-283-1
; Sequence 1, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-1
Alignment Scores:
Pred. No.: 4,738-46 Length: 524
Score: 569.00 Matches: 163
Percent Similarity: 44.03% Conservative: 106
Best local Similarity: 26.68% Mismatches: 208
Query Match: 15.97% Indels: 134
DB: 4 Gaps: 17
US-10-728-323-1 (1-2032) x US-09-424-283-1 (1-524)
QY 50 ATGAGAGGAGGAGTTTCTCCATGATGCTGTGTAGGGATCCTTGTCTGCTTCAGTT 109
Db 3 MetArgThrLysLeuSerLeuAlaIlePhePhePheLeuLeuAlaLeuPheSer--- 21
QY 110 TCTGCAAGCATGCCAAGTCATCCTTACCAGAGAAGAAACAGAACCCCTGCCCCAG 169
Db 22 AsnLeuAlaPheGlyLysCysLysGluThrGluValGluGluAspProGluLeuVal 41
QY 170 AGGTGCTCCAGAGTTGTCACAGGAACCGGAT-----GACTTGAAGCAAAAGCATGC 223
Db 42 ThrCysLysHisGlnCysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 61
QY 224 GAGTCTCGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGAC 283
Db 62 LeuGlnSerCys--- 65
QY 284 ACTGGCACCAACAACCAACGTTCCCTCCAGGGAGCGGACACGTTGGCGGCCAACCCGGA 343
Db 65 ----- 65
QY 344 GACTACGATGATACCGCGCTCAACCCGAAAGAGGAGGAGGCGGATGGGACCACT 403
Db 66 -----AspArgTyraHisArgMetLysGlnGlu----- 74
QY 404 GGCACGAGGAGCGGTGAAGAGAAAGAACCTGGAGACCAACCAAGAGAAGATTGGAGCGGA 463
Db 75 -----ArgGluLysGlnIleGlnGluThrArgGluLysLysGluGlu----- 89
QY 464 CCAAGTCATCAGCACCCACGAAATAAGGCCGAGGAGGAGGAGGAGGAGGAGGAGTGG 523
Db 90 -----GluSerArgGluArgGluGlu----- 97
```

QY 524 GGRACACAGGTAGCCATGTGAGGGAAGAACATCTCGGAACACCCCTTTCTACTTC--- 580
Db 98 -----GlnGlnGlnGlnHHisGlnGlnGlnAspGluAsnProTyrIlePheGlu 113
QY 581 CCGTCAAGGGGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGCTCTCGAG 640
Db 114 GluAspLysAspPheGluThrArgValGluThrGluGlyArgIleArgValLeuLys 133
QY 641 AGTTTGACCAAAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCCAGCAATCCATATGTGCGAGATC 700
Db 134 LysPheThrGluLysSerLysLeuLeuGlnGlyIleGluAsnPheArgLeuAlaIleLeu 153
QY 701 GAGGCCAAACCTACACTCTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGT 760
Db 154 GluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluValAlaPhePhe 173
QY 761 ATCCAGCAAGGGCAAGCCCGTACCGTAGCAATGGCAATTAACACAGAGAGCTTTAAT 820
Db 174 AsnIleLysGlyArgAlaValLeuGlyLeuValSerGluSerGluThrGluLysIleThr 193
QY 821 CTTGACGAGGGCCATGCACTCAGAAATCCCATCCGTTTTCATTTCTCATCTTTGAACCCG 880
Db 194 LeuGluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArg 213
QY 881 CATGACACACGACCTCAGAGTAGCTAAATCTCCATGCC---GTTAACACACCC 934
Db 214 AspGluAsnAspLysLeuPheLeuAlaMetLeuHisIleProValSerValSerThrPro 233
QY 935 GGCCAGTTTGGAGATTCTTCCCGGCGAGCGACCAATCATCTCTCTCTGAGGGC 994
Db 234 GlyLysPheGluGluPheAlaProGlyArgAspProGluSerValLeuSerAla 253
QY 995 TTCAGCAGGAATAGTTGGAGCGCTTCAATCGGAAATTCATGAGATACGAGGGTG 1054
Db 254 PheSerTrpAsnValLeuGlnAlaLeuGlnThrProLysGlyLysLeuGluAsnVal 273
QY 1055 CTGTTAGACAGATGTCAGAGGTGACCAAGAGAGAGGGCGAGGGCGATGGTACT 1114
Db 274 Phe----- 274
QY 1115 CGGAGTAGTGAGAACATGAAGGATGATAGTCAAAGTGTCAAAGGAGACGTTGAAGAA 1174
Db 275 -----AspGlnGlnAsnGluGlySerIlePheArgIleSerArgGluGlnValArgAla 292
QY 1175 CTTATTAGCACGCTAAATCCGCTCTCAAGAAAGGCTCC-----GAAGAAGAGGAGAT 1228
Db 293 Leu-----AlaProThrLysLysSerSerTrpTrpProPheGlyGlyGlu 307
QY 1229 ATCACCAACCAATCAACTTGAGAGAGGGCGAGCCGATCTTCTTAACAATTTGGGAG 1288
Db 308 SerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGlyTyrGlyArg 327
QY 1289 TTATTGAGGTGAAGCCA---GACAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATG 1345
Db 328 LeuThrGluValGlyProAspAspGluLysSerTrpLeuGlnArgLeuAsnLeuMet 347
QY 1346 CTCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCAAAAGGCC 1405
Db 348 LeuThrPheThrAsnIleThrGlnArgSerMetSerThrIleHisTyrAsnSerHisAla 367
QY 1406 ATGGTTATCGTCGTCAACAAGGAACTGGAAACCTTGAACCTCGTGGCTGTGAAGAAA 1465
Db 368 ThrLysIleAlaLeuValIleAspGlyArgGlyHisLeuGlnIleSerCysProHisMet 387
QY 1466 GAGCAACAACAGAGGGGACCGCGGAAGAAGAGAGGAGGAGCAAGCAAGAGAGGAGGA 1525
Db 388 SerSerArgSerSerHisSerLysHisAspLys----- 398
QY 1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTGTGAAGAGGCGATGTGTTTCATCATG 1585
Db 399 SerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPheValVal 418

QY 1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCTCGAACTCCATCTGCTT-----GGC 1639
Db 419 ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuMetIleCys 438
QY 1640 TTCGGTATCAACGCTCAAAACACACACAGAAATCTTCTTGCAGGTGATAAGGACAATGTG 1699
Db 439 PheGluValAsnAlaArgAspAsnLysPheThrPheAlaGly---LysAspAsnIle 457
QY 1700 ATAGACAGATGAGAGCAAGCAAGAGGATTTAGCATTTCTCTGGTGGGTGCAACAAGTT 1759
Db 458 ValSerSerLeuAspAsnValAlaLysGluLeuAlaPheAsnTyrProSerGluMetVal 477
QY 1760 -----GAGAACCTCATC----- 1771
Db 478 AsnGlyValPheLeuLeuGlnArgPheLeuGluArgLysLeuIleGlyArgLeuTyrHis 497
QY 1772 -----AAAAACCAAGGAATCTCACTTT 1795
Db 498 LeuProHisLysAspArgLysGluSerPhePhe 508
RESULT 12
US-09-323-195A-18
; Sequence 18, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Pertera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Picea glauca
US-09-323-195A-18
Alignment Scores:
Pred. No.: 3,57e-45 Length: 448
Score: 559.50 Matches: 134
Percent Similarity: 50.12% Conservative: 82
Best Local Similarity: 31.09% Mismatches: 178
Query Match: 15.71% Indels: 37
DB: 4 Gaps: 10
US-10-728-323-1 (1-2032) x US-09-323-195A-18 (1-448)
QY 545 AGGGAAGAAACATCTCGGAACCAACCCCTTTCTACTTCCCGTCAAGGCGGTTTAGCACCCGC 604
Db 44 ArgGluGluArgGluGluAsnProTyrValPheHisSerAspSerPheArgThrArg 63
QY 605 TACGGGAACCAAAACGGTAGGATCCGGGTCCTGAGAGGTTTGACCAAGGTCGAGGCAG 664
Db 64 AlaSerSerGluAlaGlyGluIleArgAlaLeuProAsnPheGlyGluValSerGluLeu 83
QY 665 TTTCAGAACTCCAGAAATCACCGTATTGTGCAGATCGAGGCCCAACCTTAACACTCTTGT 724
Db 84 LeuGluGlyIleArgLysPheArgValThrCysIleGluMetLysProAsnThrValMet 103
QY 725 CTTCCCAACGACGCTGATCTGATAACATCTTGTATTATCCAGCAAGGGCAAGCCACCGTG 784
Db 104 LeuProHisTyrIleAspAlaThrIleLeuTyrValThrArgGlyArgGlyTyrIle 123
QY 785 ACCGTAGCAAAATGGCAATTAACAAAGAGCTTTAATCTTGACGAGGGCCATGCACCTCAGA 844
Db 124 AlaTyrValHisGlnAsnGluLeuValLysArgLysLeuGluGluGlyAspValPheGly 143
QY 845 ATCCCATCCGGTTTCATTTCTTACATCTTGTGAACCCGATGACCAACGAGACCTCAGAGTA 904
DB: 4 Gaps: 10

```
Db 144 ValProSerGlyHisThrPheTyrLeuValAsnAsnAspHisAsnThrLeuArgIle 163
QY 905 GCTAAATCTCCATGCGGTAAACACA--CCGGCCAGTTTGAGGATTTCTCCGGCG 961
Db 164 AlaSerLeuValArgProValSerThrValArgGlyGluTyrGlnProPheTyrValAla 183
QY 962 AGCAGCCGAGACCAATCTCTACTTGCAGGGCTTCAGCAGGAATACGTTGGAGCGCGC 1021
Db 184 GlyGlyArgAsnProGlnThrValTyrSerAlaPheSerAspValLeuGluAlaAla 203
QY 1022 TTCATGCGGAATTCATGAGATACGAGGGTCTGTTAGAGAGAAATCCAGGAGGTAG 1081
Db 204 PheAsnThrAsnValGlnGlnLeuGluArgIlePhe-----GlyGly---- 217
QY 1082 CAAGAGGAGAGGGCAGAGGCGATGGAGTACTCGGAGTAGTGAGAACAAATGAAGGAGTG 1141
Db 218 -----HisLysSerGlyVal 222
QY 1142 ATAGTCAAAAGTGTCAAAGGAGCAGTTGAAGAACTTACTAAGCACGCTAAATCCGTTCA 1201
Db 223 IleIleHisAlaAsnGluGlnIleArgGluMetMetArg-----LysArgGlyPhe 240
QY 1202 AAGAAAGGTCGGAAGAGAGGAGATATCACCAACCCAAATCAACTTGAGAGAGGCGAG 1261
Db 241 SerAlaGlySerMetSerAlaProGluHisProLysProPheAsnLeuArgAsnGlnLys 260
QY 1262 CCGGATCTTTCTAACAACTTTGGGAAGTTATTTGAGGTGAGCCAGCACAGAGAACCCC 1321
Db 261 ProAspPheGluAsnGluAsnGlyArg---PheThrIleAlaGlyProLysAsnTyrPro 279
QY 1322 CAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGGAGCTTTGATG 1381
Db 280 PheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeuAsnProGlySerMetThr 299
QY 1382 CTCGCCACACTTCAACTCAAGGCCATGTTATGCTGCTGCTGCACAAAGAACTGGAAC 1441
Db 300 AlaProSerLeuAsnSerLysSerThrSerIleGlyIleValThrAsnGlyGluGlyArg 319
QY 1442 CTTGAACTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGCGGGGAGAGAGGAG 1501
Db 320 IleGluMetAlaCysProHisLeuGlyGlnHisGlyTyrPheSerProArgGluArgGly 339
QY 1502 GACGAAGACGAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTG 1561
Db 340 AspGlnAsp-----IleThrTyrGlnArgValTrpAlaLysLeu 352
QY 1562 AAGGAAGGGAGTGTTCATCATGCCAGAGCTCATCCAGTAGCCCATCAACGCTTCC--- 1618
Db 353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr 372
QY 1619 ---TCCGAACTCCATCTGCTTGGCTTCGGTATCAAGCTGAAACACACAGAAATCTTC 1675
Db 373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe 392
QY 1676 CTTGCAAGGTGATAAGGACAAATGTGATGACACAGATAGAACGCAAGGAGGATTTAGCA 1735
Db 393 LeuAlaGly---LysAsnAsnValLeuAsnThrLeuGluArgGluIleArgGlnLeuSer 411
QY 1736 TTC---CCTGGTGGGTGAACAAGTTGAGAGCTCATCAAAACACAGAAAGAAATCTCAC 1792
Db 412 PheAsnValProArgGlyGluGluIleGluValLeuGlnAlaGlnLysAsp---Gln 430
QY 1793 TTGTGAGTGCTGCTCCTCAATCTCAATCTCAA 1825
Db 431 ValIleLeuArgGlyProGlnArgArgSerArg 441
```

RESULT 13

US-09-424-283-4

; Sequence 4, Application US/09424283

; Patent No. 6437219

; GENERAL INFORMATION:

; APPLICANT: Grimes, et al.

; TITLE OF INVENTION: Sucrose binding proteins

```
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-4
Alignment Scores: 7,75e-41 Length: 409
Score: 514.50 Matches: 133
Percent Similarity: 42.72% Conservative: 87
Best Local Similarity: 25.83% Mismatches: 168
Query Match: 14.44% Indels: 127
DB: 4 Gaps: 10
```

US-10-728-323-1 (1-2032) x US-09-424-283-4 (1-409)

```
QY 71 CTGATGCTGTTGTAGGATCCTTCTGCTGCTTCTGCAACGCATGCCAAGTCA 130
Db 16 LeuLeuAlaLeuIleSerAsnLeuAlaLeuGlyLysLeuLysGluThrGluValGlu--- 34
QY 131 TCACCTTACCAGAAACAGAACCCCTCGGCCAGAGGTGCTCCAGAGTTGTCAA 190
Db 35 -----GluAspProGluLeuValThrCysLysHisGlnCysGln 47
QY 191 CAGGAACCGGAT-----GACTTGAAGCAAAAGGCATCGGAGTCTCGTCGACCAAGCTC 244
Db 48 GlnGlnArgGlnTyrThrGluSerAspLysArgThrCysLeuGlnGlnCysAspSerMet 67
QY 245 GAGTATGATCCTCGTGTGTCTATGATCCTCGAGACACACTGGCACCAACCAACGCT 304
Db 68 LysGlnGlu----- 70
QY 305 TCCCTCCAGGGAGCGGACACAGTGGCGGCCAAACCCGAGAGACTAGATGATGACCGCGGT 364
Db 70 ----- 70
QY 365 CAACCCCGAAGAGAGAGAGGCCGATGGGACCCAGCTGGACCGAGGGAGCGTGAAGA 424
Db 71 -----ArgGluLysGlnVal 75
QY 425 GAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGGCGACCAAGTCAATCAGCAGCCACGG 484
Db 76 GluGluGluThrArgGluLysGluGlu-----HisGlnGluGln--- 89
QY 485 AAATAAGCCCGAAGGAAGAGAGAGAAACAAGAGTGGGGAACACCAAGGTAGCCATGTG 544
Db 89 ----- 89
QY 545 AGGGAAGAAACATCTCGGAAACAACCTTTCTACTTC---CCGTCAAGCGCGGTTTAGCACC 601
Db 90 HisGluGluGluGluAspGluAsnProTyrValPheGluGluAspLysAspPheSerThr 109
QY 602 CGCTACGGGAACCAACGAGGTAGGATCCGGTCTCTGACAGAGTTTGACCAAGGTCAAGG 661
Db 110 ArgValGluThrGluGlyGlySerIleArgValLeuLysLysPheThrGluLysSerLys 129
QY 662 CAGTTTTCAGATCTCCAGATCAACCGTATTGTGAGATCGAGGCGCAAACTTAACACTCTT 721
Db 130 LeuLeuGlnGlyIleGluAsnPheArgLeuAlaIleLeuGluAlaArgAlaHisThrPhe 149
QY 722 GTTCTTCCAGACGCTGATGCTGATAACATCTCTGTTATCTCCAGCAAGGCAACCCACC 781
Db 150 ValSerProArgHisPheAspSerGluValValLeuPheAsnIleLysGlyArgAlaVal 169
```



```
Db      274 Phe----- 274
QY      1115 CGAGTAGTGAGAACATGAAGAGGTGATAGTCAAAAGTGTCAAAGGAGCACGTTGAAGAA 1174
Db      275 -----AspGlnGlnAsnGluSerIlePheArgIleSerArgGluGlnValArgAla 292
QY      1175 CTTACTAAGCACGCTAATCCGCTCAAGAAAGGCTCC-----GAAGAAGAGGGAGAT 1228
Db      293 Leu-----AlaProThrLysSerSerTyrPheGlyGlyGlu 307
QY      1229 ATCAACCAACCAATCAACTTGAGAGAAGCGGAGCGATCTTTCTAACCAACTTTGGGAAG 1288
Db      308 SerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGlyTyrGlyArg 327
QY      1289 TTATTTGAGGTGAAGCCA---GACAAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATG 1345
Db      328 LeuThrGluValGlyProAspAspGluLysSerTyrLeuGlnArgLeuAsnLeuMet 347
QY      1346 CTCACCTGTGTAGAGATCAAGAAGAGAGCTTTGATGCTCCACACTTCACTCAAAAGGCC 1405
Db      348 LeuThrPheThrAsnIleThrGlnArgSerMetSerThrIleHisTyrAsnSerHisAla 367
QY      1406 ATGTTTATCGTCTGCTCAACAAAGGAACCTTGGAACCTTGAACCTCGTGGCTGTAAAGAAA 1465
Db      368 ThrLysIleAlaLeuValIleAspGlyArgGlyHisLeuGlnIleSerCysProHisMet 387
QY      1466 GAGCAACAACAGAGGGGACCGCGGGAAGAGAGAGGAGGACGAGACGAAAGACGAAAGAGAGGAGGA 1525
Db      388 SerSerArgSerSerHisSerLysHisAspLys----- 398
QY      1526 AGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTCGAAGAGGCGGATGTGTTTCATCG 1585
Db      399 SerSerProSerTyrHisArgIleSerSerAspLeuLysProGlyMetValPheValVal 418
QY      1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACTCCATCTGCTT-----GGC 1639
Db      419 ProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuMetIleCys 438
QY      1640 TTCGGTATCAACGCT 1654
Db      439 PheGluValAsnAla 443
```

Search completed: August 24, 2005, 10:09:09
Job time : 62.1903 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 24, 2005, 10:05:17 ; Search time 131.012 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3562
Sequence: 1 ataatacatatatattcatc.....cgtttgctgctgtttctcc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DRV=xlh
-Q=/cgn2.1/USPTO spool/US10728323/runat_23082005_124357_29323/app_query.fasta_1.4757
-DB=Published Applications AA -OEMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -NODES=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10728323 @CGN 1 1 221 @runat 23082005 124357_29323
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCAPOP=6 -FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2.6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2: /cgn2.6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2.6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2.6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2.6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2.6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2.6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8: /cgn2.6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2.6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgn2.6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgn2.6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgn2.6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13: /cgn2.6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgn2.6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgn2.6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgn2.6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgn2.6/ptodata/2/pubpaa/US10E_PUBCOMB.pcp.*
18: /cgn2.6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
19: /cgn2.6/ptodata/2/pubpaa/US11A_PUBCOMB.pcp.*
20: /cgn2.6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
21: /cgn2.6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
22: /cgn2.6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3286	92.3	626	10	US-09-847-208-28	Sequence 28, Appl
2	3286	92.3	626	14	US-10-228-806-2	Sequence 2, Appl
3	3286	92.3	626	15	US-10-100-303A-7	Sequence 7, Appl
4	3286	92.3	626	15	US-10-245-871-10	Sequence 10, Appl
5	3286	92.3	626	15	US-10-253-286-10	Sequence 10, Appl
6	3286	92.3	626	16	US-10-809-689-95	Sequence 95, Appl
7	3286	92.3	626	17	US-10-899-551-2	Sequence 2, Appl
8	3192	89.6	635	17	US-10-899-551-53	Sequence 53, Appl
9	3052	85.4	634	9	US-09-731-221-78	Sequence 78, Appl
10	3041	85.4	614	9	US-09-331-631A-21	Sequence 21, Appl
11	3041	85.4	614	10	US-09-847-208-27	Sequence 27, Appl
12	3041	85.4	614	14	US-10-147-095-21	Sequence 21, Appl
13	3041	85.4	614	15	US-10-100-303A-8	Sequence 8, Appl
14	1394	39.1	268	15	US-10-100-303A-55	Sequence 55, Appl
15	1282	36.0	605	15	US-10-424-599-153195	Sequence 153195
16	1281	36.0	605	9	US-09-331-631A-25	Sequence 25, Appl
17	1281	36.0	605	14	US-10-147-095-25	Sequence 25, Appl
18	1275.5	35.8	623	15	US-10-424-599-153206	Sequence 153206
19	1274.5	35.8	623	16	US-10-739-930-8840	Sequence 8840, Ap
20	1268	35.6	605	15	US-10-100-303A-110	Sequence 110, App
21	1201.5	33.7	584	15	US-10-424-599-260105	Sequence 260105
22	1157	32.5	417	15	US-10-245-227B-14	Sequence 14, Appl
23	1157	32.5	425	15	US-10-245-227B-12	Sequence 12, Appl
24	1156	32.5	439	15	US-10-245-227B-1	Sequence 1, Appl
25	1117.5	31.4	390	15	US-10-245-227B-2	Sequence 2, Appl
26	874	24.5	324	15	US-10-425-114-44408	Sequence 44408, A
27	867	24.3	344	15	US-10-425-114-51703	Sequence 51703, A
28	865.5	24.3	666	9	US-09-331-631A-3	Sequence 3, Appl
29	865.5	24.3	666	14	US-10-147-095-3	Sequence 3, Appl
30	863.5	24.2	625	9	US-09-331-631A-5	Sequence 5, Appl
31	863.5	24.2	625	14	US-10-147-095-5	Sequence 5, Appl
32	849.5	23.8	666	9	US-09-331-631A-1	Sequence 1, Appl
33	849.5	23.8	666	14	US-10-147-095-1	Sequence 1, Appl
34	828	23.2	316	15	US-10-425-114-43905	Sequence 43905, A
35	828	23.2	316	15	US-10-425-114-49262	Sequence 49262, A
36	827	23.2	315	15	US-10-425-114-43971	Sequence 43971, A
37	816	22.9	301	15	US-10-425-114-49373	Sequence 49373, A
38	791	22.2	590	9	US-09-331-631A-8	Sequence 8, Appl
39	791	22.2	590	14	US-10-147-095-8	Sequence 8, Appl
40	790	22.2	300	15	US-10-425-114-43597	Sequence 43597, A
41	788	22.1	296	15	US-10-425-114-43984	Sequence 43984, A
42	770	21.6	291	15	US-10-425-114-51411	Sequence 51411, A
43	761	21.4	149	15	US-10-100-303A-54	Sequence 54, Appl
44	737.5	20.7	525	9	US-09-331-631A-7	Sequence 7, Appl
45	737.5	20.7	525	14	US-10-147-095-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-847-208-28
; Sequence 28, Application US/09847208
; Publication NO. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-28

Alignment Scores:									
Pred. No.:	2.4e-261	Length:	626						
Score:	3286.00	Matches:	626						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	92.25%	Indels:	0						
DB:	10	Gaps:	0						
US-10-728-323-1 (1-2032) x US-09-847-208-28 (1-626)									
QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTGTAGGGATCCTTGCTCCTGCTTCAGTT	109						
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20						
QY	110	TCTGCAACGCATGCCAAGTCATCCTTACCAAGAAGAAAACAGAGAACCCCTGCGCCACG	169						
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40						
QY	170	AGGTGCCTCCAGATTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCATCGAGTCT	229						
Db	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60						
QY	230	CGCTCACCAAGCTCGAGTATGATCCTCGTTCTGTCTATGATCCTCGAGGACACACTGCG	289						
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80						
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGCTGGCGGCAACCCCGAGACTAC	349						
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100						
QY	350	GATGATGACCGCGTCAACCCGAGAGAGGAGGAGGCCGATGGGACCACTGGAACCG	409						
Db	101	AspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro	120						
QY	410	AGGGAGCGCTGAAGAAGAAGACTGCGAGACAACCAAGAGAAGATTGGAGGCGCAACAGT	469						
Db	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	140						
QY	470	CATCAGACGCCACGAAATTAAGGCCGAGAGAGAGAGAGAAACAAGTGGGGAACA	529						
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr	160						
QY	530	CCAGTAGCCATGTGAGGAGAGAAACATCTCGGAACAACCTTTCTACTTCCGTCAGG	589						
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180						
QY	590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGTTTGAC	649						
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200						
QY	650	CAAGGTCAAGCGAGTTTCAGAAATCCAGAAATCACCGTATGTGTCAGATCAGAGCCAAA	709						
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220						
QY	710	CCTAACACTCTGTTCCTCCAGACGCGTGTCTGATTAACATCCTTGTATTCCAGCA	769						
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240						
QY	770	GGGCAAGCCACCGTACCGTAATGGCAATTAACAGAAAGAGCTTTTAATCTTCACGAG	829						
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260						
QY	830	GGCCATGCACTCAGAAATCCCATCCGGTTTCATTCTCTACATCTTGAACCGCCATGACAC	889						
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280						
QY	890	CAGAACCTCAGAGTAGTAAATCTCCATGCCGTTTAAACACACCGCGCCAGTTTCAGAT	949						
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300						
QY	950	TTCTTCCGCGGAGCGAGCCGAGACCAATCATCTACTTTCGAGGGCTTCAGCAGGAATACG	1009						

```
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228, 806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-2

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 14 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-228-806-2 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCATGATGCTGTCTAGGGATCCTTGTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TTGCAACGATGCGCAAGTCATCACCTTACCAGAGAAAACAGAGAAACCCCTGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGATTGTCACAGGAACCGGATGATCTCAAGCAAAAGCATGCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGGTGCACCAAGCTCCAGTATGATCTCTGTGTGTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluLysArgProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACAGTTCCCTCAGGGAGCGGACACAGTGGCGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGCGCGCTCAACCCGAGAGAGGAGGCGGATGGGGACCACTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrProAlaGlyPro 120
QY 410 AGGGAGCGTGAAGAGAGAAGACTCGAGACAACCAAGAGAAGATTGGAGGCGCAACAGT 469
DB 121 ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGACCCACGGAATAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGlnIleGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAAGAAACATCTCGGAACAACCTTCTACTTCCGTCAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAACACGGTAGGATCCGGTCTCTGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCGAGTTTCAGAAATCCAGAAATCACCGTATTGTGCGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCAGCTGATGCTGATAAATCTCTTCTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAATGGCAATTAACAGAAAGAGCTTTTAACTCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
```

```
QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTTACATCTTGAACCGCCCATGACAA 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCCAGTTTGGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCGGAGACCAATCATCTTCTAGGGCTTTCAGCAGGATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGAGCGCCTTCAATGCGGAATTCATGAGATACGAGGGTGTCTGTAGAGAGAA 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGCGGATGAGTACTCGGAGTAGTGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrPseThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGCACGTTTGAAGAACTTACTAAGCACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAGAGAGGAGATATCACCACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGCGGCGCCGATCTTCTAAACAACCTTTGGAGAGTATTATTGAGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspPheSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAAGACCCCGACCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTGTCTCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTGTGTGCTGTAAAGAAAGAGCAACACAGAGGGAGCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 CAAAGAGAGGAGGAGCAAGAGAGAGAGAGGAGGAGGAGTAAACAGAGAGTGTCTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTGAAGAGAGGCGATGTGTTCATCATGCTCCAGCAGCTCATCCAGTACCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCATCTCGCTTCCGTTTCGATCAACGCTGAAACCAACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTTCAGGTTGATAAGGACAAATGTGTAGACCAAGATAGAGAGCAAGCGAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAGAGTTCATCAAAAACACAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCCTCAATCTCAATCTCAATCTCGTCTCTCTCTCTCTCTCTCTCT 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB 601 SerProGluLysGluAspGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 620
```


QY	1910	ATTTGAAGCGTTTAAAC	1927	
Db	621	IleLeuLysAlaPheAsn	626	
RESULT 3				
US-10-100-303A-7				
; Sequence 7, Application US/10100303A				
; Publication No. US20030202980A1				
; GENERAL INFORMATION:				
; APPLICANT: Caplan, et al.				
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction				
; TITLE OF INVENTION: to Allergy				
; FILE REFERENCE: 2002834-0166				
; CURRENT APPLICATION NUMBER: US/10/100,303A				
; CURRENT FILING DATE: 2002-03-18				
; NUMBER OF SEQ ID NOS: 138				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 7				
; TYPE: PRT				
; LENGTH: 626				
; ORGANISM: Atachis hypogaea, Prot/Nucleo-Ara h 1				
US-10-100-303A-7				
Alignment Scores:				
Pred. No.:	2.4e-261	Length:	626	
Score:	3286.00	Matches:	626	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	92.25%	Indels:	0	
DB:	15	Gaps:	0	
US-10-728-323-1 (1-2032) x US-10-100-303A-7 (1-626)				
QY	50	ATGAGAGGAGGTTTCTCCACTGATGCTGTGCTAGGGATCCTTGCTCGTCTTCAGTT	109	
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuGlyIleuValLeuAlaSerVal	20	
QY	110	TCTGCAACGCATGCCAAGTCATCACCTTACCAGAAAGAAAACAGAGAACCCCTTCGCCCCAG	169	
Db	21	SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40	
QY	170	AGTGCTCCAGATTGTCACAGGACCGGATGATCTGAAGCAAAAGCATCGAGTCT	229	
Db	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60	
QY	230	CGCTGCACCAAGCTCGAGTATCATCTCGTCTGCTATGATCCTCGAGGACACACTGCG	289	
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80	
QY	290	ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGGCCGCCCAACCCCGAGACTAC	349	
Db	81	ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100	
QY	350	GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGGACCACTGGACCG	409	
Db	101	AspAspAspArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro	120	
QY	410	AGGGAGCGTGAAGAGAGAGACTGGGACACCAAGAGAGATTGGAGGCGCAACT	469	
Db	121	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer	140	
QY	470	CATCAGCAGCCAGGAAATAAGGCCGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG	529	
Db	141	HisGlnGlnProArgLysIleArgProGlyArgGlyGluGlnGluTrpGlyThr	160	
QY	530	CCAGTAGCCATGTGAGGAGAGAAACATCTCGGAACAACCCCTTCTACTCCCGCTCAAGG	589	
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180	
QY	590	CGTTTAGCCCGCTTACGGGACCAAAACGCTAGGATCCGGGTCTCGCAGAGGTTTGAC	649	
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValIleuGlnArgPheAsp	200	

QY	650	CAAAGGTCAAGCAGTCTTTCAGAAATCCAGAAATCACCTATTGTGCAGATCGAGCCAAA	709	
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220	
QY	710	CCTAACACTCTGTTCITTCCTCCAGCAGCCTGATGCTGATACATCCTTGTATCCAGCNA	769	
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240	
QY	770	GGGCAAGCCACCGTCAGCTAGCAATGCAATACAGAAAGAGCTTTAATCTTCACGAG	829	
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260	
QY	830	GGCCATGCACCTCAGAAATCCCATCCGTTTTCATCTTCTACATCTTGAACCGCATCACAC	889	
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280	
QY	890	CAGAACCTCAGAGTAGCTAAATCTCCATGCCCTTAACACACCCCGGCAGTTTACGAT	949	
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300	
QY	950	TTCTTCCCGCGAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATACG	1009	
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320	
QY	1010	TTGGAGCGCGCTTCAATGCCGAATTCATGAGATACGAGGGTCTCTTAAAGAGAGAT	1069	
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340	
QY	1070	GCAGAGGTGACCAAGAGAGAGAGGCGAGGGCGATCGAGTACTCGGAGTAGTCAGAAC	1129	
Db	341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360	
QY	1130	AATGAAGAGGTGATAGTCAAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCACGCT	1189	
Db	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380	
QY	1190	AAATCCGTCTCAAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCCCAATCAACTTG	1249	
Db	381	LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu	400	
QY	1250	AGAGNAGCGAGCCGATCTTCTTAACAACCTTGGGAAGTTATTGAGGTGAGGAGCCAGAC	1309	
Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420	
QY	1310	AAGAAGAAACCCCGCTTCAGGACCTGCACATGATGCTCACCTGTGTAGAGATCAAGAA	1369	
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluLleLysGlu	440	
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCGTCAACAAA	1429	
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460	
QY	1430	GGAACTGGAACCTTGAACCTCGTGGCTGTAGAAAGAGCAACACACAGGGGCGCGG	1489	
Db	461	GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg	480	
QY	1490	GAAGAAGAGGAGGAGCAAGCAAGAGAGGAGGAGTAACACAGAGCTGCTAGGTAC	1549	
Db	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500	
QY	1550	ACAGCAGGTTGAAGGAAGCGATGTGTTTCATCATGCGCAGAGCTCATCCAGTACCCATC	1609	
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520	
QY	1610	AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAACCAACACACA	1669	
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540	
QY	1670	ATCTTCTTCAGAGGTGATAAGCAATGTGTATAGACCATAGAGAACCAAGCAGGAGAT	1729	
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560	
QY	1730	TTAGCATTTCCCTGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT	1789	


```
Db 481 GluGluGluGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGAACGCGATGTTCATCATCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuGlyGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCCATCTGCTTGCTTTCGCTATCAACGCTGAAACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
QY 1670 ATCTTCTTGCAGTGATGAAGACATGTGATAGCAGATAGAGAACGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspIleAspValIleAspGlnIleGluGlyGlnAlaIleAsp 560
QY 1730 TTAGCATTCCTCGGTCGAGTGAACAGTTGAGAGCTCATCAAAACACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluGlyLeuIleAsnGlnIleGlySer 580
QY 1790 CACTTTGTGAGTCTGCTCTCAATCTCAATCTCCTGCTGCTCTCTCCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluGlyGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGAAACCAAGAGGAGGAGTCCATCTCTTCA 1909
Db 601 SerProGluGlyGluAspGlnGluGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTAAAC 1927
Db 621 IleLeuGlyAlaPheAsn 626

RESULT 5
US-10-253-286-10
; Sequence 10, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: I1-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-10

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-253-286-10 (1-626)
QY 50 ATGAGAGGAGGTTCCTCCATGATGCTGTGCTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValIleuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAGAGAAACACAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaIleSerSerProTyrGlnIleGlySerThrGluAsnProCysAlaGln 40
QY 170 AGTGCCCTCCAGAGTTGTCAACAGGAACCGGATGATCTTGAAGCAATCTTAAAGCACCTT 229
```

```
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuGlySerGlnLysAlaCysGluSer 60
QY 230 CGCTCACCACAGCTCGAGTATGATCCTCGTGTGTCGTATGATCCTCGAGGACACACTCGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGCGGACACGCTGGCCGCCAACCCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCCGTCFAACCCCGAAGAGAGAGAGCCGATGGGACCCAGCTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgGlyGlnProAlaGlyPro 120
QY 410 AGGAGCGCTGAAGAGAGAGAGACTGGAGAGACCAACAGAGAGATTGAGGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGGAATAAGGCCCGAAGAGAGAGAGAGAACAAAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgGlyIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGTACCATGTGAGGAGAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAGAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGNAACCAAAACGTTAGGATCCGGTCTCTCAGAGGTTTCAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGluArgPheAsp 200
QY 650 CAAAGGTCACAGCGAGTTTCAGAAATCTCCAGAAATCACCTATTGTGCAGATCAGAGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTGTTCTTCCCAAGCACGCTGATCTGATGATACATCTTGTATTCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCACCGTAGCAATGCGCAATACAGAGAGCTTTAACTTTCACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACCTCAGAAATCCCATCCGTTTCTATCTTCTATCATCTTTGAACCGCATCACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCGCAGTTTGAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTGTGAGGCTTCAGAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCGCTTCAATGCCGAATTCATAGATACACGAGGCTGCTCTTGAAGAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGAGGTGAGCAAGAGAGAGAGGCGAGGCGATGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATGATCAAAAGTGTCAAAGGAGACGTTGAAGAACTTACTTAAGCACCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AAATCCGCTCTCAAAAGAGGCTCCGAAAGAGAGGAGATATCACCAACCCCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGCGGAGCCCGATCTTTCTAAACACTTGGGAAGTATTGAGGTGAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
```

QY 1310 AACGAAGACCCAGCTTCAGGACCTGACATGCTCCTGCTAGAGATCAAGAA 1369
 DB 421 LysLeuAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
 QY 1370 GGAAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCGTCACAAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
 QY 1430 GGAATCGGAACCTTGAATCTGCTGGCTTAAGAAAAGACAAACAGAGGGGACGGCG 1489
 DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
 QY 1490 GAAGAAGAGGAGACCAAGACGAGAGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGCTAC 1549
 DB 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCAGGTTGAAGGAGCGGATGTGTCATCATGCCAGCAGTCATCCAGTAGCCATC 1609
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
 QY 1610 AACGCTTCCTCGAATCCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTTCTTCAGGTGATTAAGGACATGTCATAGACAGATAGACAGAGCAAGCAAGGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAGAAAGAACTT 1789
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCCGTCGTCCTCGAGAAAGAG 1849
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCACTCTTCA 1909
 DB 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGCTTTTAAC 1927
 DB 621 IleLeuLysAlaPheAsn 626

RESULT 6

; Sequence 95, Application US/10809689
 ; Publication No. US20040265342A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eric Potter Clarkson
 ; TITLE OF INVENTION: Methods and compositions for desensitisation
 ; FILE REFERENCE: 5538/1010
 ; CURRENT APPLICATION NUMBER: US/10/809,689
 ; CURRENT FILING DATE: 2004-03-25
 ; PRIOR APPLICATION NUMBER: PCT/GB99/00080
 ; PRIOR FILING DATE: 1999-01-11
 ; PRIOR APPLICATION NUMBER: GB/9800445.0
 ; PRIOR FILING DATE: 1998-01-09
 ; PRIOR APPLICATION NUMBER: GB/9820474.6
 ; PRIOR FILING DATE: 1998-09-21
 ; NUMBER OF SEQ ID NOS: 124
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 95
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: *Arachis hypogaea*
 US-10-809-689-95

Alignment Scores:

Pred. No.: 2,4e-261 Length: 626
 Score: 3286.00 Matches: 626
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.25% Indels: 0
 DB: 16 Gaps: 0
 US-10-728-323-1 (1-2032) x US-10-809-689-95 (1-626)
 QY 50 ATGAGAGGGAGGTTTCTCCACTGATGCTGTGTAGGGATCCTGTCTGCTTCCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACCATGCGCAAGTCATCAGCTTACACAGAGAAAAACAGAGAACCCCTGCGCCCCAG 169
 DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
 QY 170 AGTGCTCCAGAGTTGTCAACAGAGAACCGGATAGCTTGAAGCAAAAGAGCATCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
 QY 230 CGCTGCACCAAGCTCGAGTATGATCCTGTTGCTGTCTATGATCCTCGAGGACACATGCG 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACAGTGGCGCCCAACCCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGTCAACCCCAAGAGAGAGAGAGGAGGCGGATGGGACCAAGTGGACCG 409
 DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
 QY 410 AGGGAGCGTGAAGAGAGAGAACTGGAGACAAACCAAGAGAGAGATGGAGGCGACCAACT 469
 DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
 QY 470 CATCAGACGCCACGGAATAAGGCCGCAAGAGAGAGAGAGAGAAACAGAGTGGGGAACA 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 160
 QY 530 CAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTTCTCCGCTCAAGG 589
 DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTACACCCGCTACGGGAACCAAAACGCTAGGATCCGGGTCTCTGAGAGGTTTGAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCCAGGATTCAGATCCAGAGGCGCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAACACTCTTTGTTTCTTCCCAAGCACGCTGATGCTGTAAACATCTTGTATTCCAGCAA 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTCACCGTAGCAATGCGCATACAGAAAGCTTTTATCTTGACGAG 829
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCCATGCACTCAGAAATCCCATCCGTTTCAATTTCTTCTACATCTTGAACCGCATGACAAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
 QY 890 CAGAACTCTCAGAGTAGCTAAATCTCCATGCGCCGTTTAAACACACCCGCGCAGTTTGAAGAT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTTCCCGCGGAGCGGAGCGGAGACCAATCATCTTCTACTTTCAGGCTTCAGCAGGATACG 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGGCGCGCTTCAATGCGGAAATTCATGATGATACGAGGGGTGCTGTTAGAGAGAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340

```
QY 1070 GCAGGAGGTGACCAAGAGGAGAGGCGCAGAGCGATCGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGlnGluGlnArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTGATAGTCAAAAGTGTCAAAGGACGACGTTGAAGAACTTACTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIleValIle 380
QY 1190 AATCGCTCTCAAGAAAGCGCTCCGAGAGAGGAGGAGATATCACCACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAGGCGAGCCCGATCTTCTTAACAACTTTGGGAAGTATTATTGAGTGGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 420
QY 1310 AGAGAAGACCCAGCTTCAGGACCTCGAGACATGATGCTCCTCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTCAACTCAAGGCCATGTTATCGTCTGCTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValValValVal 460
QY 1430 GGAACCTGGAAACCTTGAACTCGTGGCTGTAAGAAAGAGCAACACAGAGGGGAGCGGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValAlaValArgLysGluGlnGlnArgGlyArg 480
QY 1490 GAAGAAGAGGAGACCAAGACCAAGAGAGGAGGAAAGTAAACAGAGAGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGGAAGAGCGGATGTTTCATCATGCCAGCGACTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATGAAGACAAATGTGATAGACCAAGATAGAGAAGCAAGAGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGGTGCAACAGTTGAGAGCTCATCAAAACCAAGAGCAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGCTCTCTCCTCAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGAGGAGGTCCTCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCGTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
```

RESULT 7

```
US-10-899-551-2
; Sequence 2. Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
```

```
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 626
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-2

Alignment Scores:
Pred. No.: 2,4e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.25% Indels: 0
DB: 17 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-899-551-2 (1-626)
QY 50 ATGAGAGGAGGGTTCCTCCACTGATGCTGTGTCTAGGATCCTTGTCTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTTCATCCTTACCAGAGAAAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCCCTCAGAGTTGTCAACAGGAACCGGATGCTTGAAGCAAAAAGGCATGCCAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCTCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluLysAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGGCGACACGTGGCCGCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409
Db 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGAGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469
Db 121 ArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 140
QY 470 CATCAGACGCCACGAAATAAGGCCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTACACCCGCTACGGGAACCAAAACCGTAGTAGTCCGGTCTCTCAGAGGTTTCAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGAAATCCCGTATGTGTGAGATCAGAGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCACGCTGATGCTGATACATCTCTTGTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTCACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACCTCAGAAATCCCATCCGGTTTCATTTCTTACATCTTTGAACCGGCATGACAAC 889
```

```
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrlleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATCCCGTTAAACACACCCGCGCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaIysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCAGACCATCATCTCTTTCAGGGCTTCAGCAGGATACG 1009
Db 301 PhePheProAlaSerArgAspGlnSerSerTyrlleGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGCGAATTCATAGATACGGAGGCTGCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGATGAGGATGAGTACTCGGATGATGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAGGTGTCAAAGGACGCTTGAAGACTTACTAAGCAGCCT 1189
Db 361 AsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIle 380
QY 1190 AATTCGCTCTCAAGAAAGGCTCCGAAAGAGAGGAGATATCAACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGAGCCGATCTTTCTAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGCTTCAGGACCTCGACATGCTACCTGCTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGCCATGTTATCTGTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGTGTTAAGAAAGAGCAACACAGAGGGAGCGCG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGACCAAGACGAGAGAGAGGAGGAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTGAAGGAAGCGGATGTTCATCATCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCCATCTGCTTGGCTTCGGTATCAAGCTGAAACCAACCAAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTGATTAAGGCAATGTCATAGACCATAGACAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTGGGTGGAACAAAGTTGAGAAGCTCATCAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCCTCAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAAAACCAAGGAGGAGGCTCACTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
```

```
Db 621 IleLeuLysAlaPheAsn 626
RESULT 8
US-10-899-551-53
; Sequence 53, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 635
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-53
Alignment Scores:
Pred. No.: 1,33e-253 Length: 635
Score: 3192.00 Matches: 605
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.83% Mismatches: 0
Query Match: 89,61% Indels: 0
Gaps: 17
US-10-728-323-1 (1-2032) x US-10-899-551-53 (1-635)
QY 110 TCTGCAACCATCGCAAGTCATCACCTTACCAGAAAGAAAAACAGAGAACCCCTGCGCCAG 169
Db 17 SerSerThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 36
QY 170 AGTGCTCCAGAGTTGTCAAACAGAAACCGGATGACTTGAAGCAAAAGCATGCGAGTCT 229
Db 37 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 56
QY 230 CCCTCACCAAGCTCGAGTATGATCCTCGTGTGCTATGATCCTCGAGACACACTGCG 289
Db 57 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 76
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACGTCGCGCCGCAACCCCGAGACTAC 349
Db 77 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlnProGlyAspTyr 96
QY 350 GATGATGACCCGCTCAACCCCGAAGAGAGAGGAGGCGCGATGGGACCACTGGACCG 409
Db 97 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 116
QY 410 AGGGAGCTGGAAGAGAGAGACCTGGAGACAAACCAAGAGAGATTGGAGGCGCAACT 469
Db 117 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 136
QY 470 CATCAGACGCCACGAAATAAGCGCCGCAAGCAAGAGAGAGAGAAACAGAGTGGGGAACA 529
Db 137 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 156
QY 530 CCAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCGTCGCAAG 589
Db 157 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 176
QY 590 CGGTTTAGCACCCGCTACGGNACCAAAACGCTAGATCCGGTCTCTGCAGAGGTTTGAC 649
Db 177 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 196
QY 650 CAAAGGTCAAGGCAAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCAGGCGCAA 709
Db 197 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 216
```


QY	710	CCTAACACTCTTGTTCTTCCCAAGCAGCGCTGATGCTGATATAACATCCTTGTTTATCCAGCAA	769
Db	217	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGln	236
QY	770	GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	237	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	256
QY	830	GGCCATGCACCTCAGAATCCCATCCGGTTTCAATTTCTCATCTTTGAACCGCCATGACAAC	889
Db	257	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrlleLeuAsnArgHisAspAsn	276
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCCGCTTAAACACCCCGCGCAGTTTGAGGAT	949
Db	277	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	296
QY	950	TTCTTTCCGGCGAGCAGCGCAGAGCAATCATCTCTACTTGCAAGGCTTCAGCAGGAATAACG	1009
Db	297	PhePheProAlaSerSerArgAspGlnSerSerTyrlleGlnGlyPheSerArgAsnThr	316
QY	1010	TTGAGGCGCGCTTCAATGCGGAATTCATAGATACGAGGGTCTGTTAGAGAGAAT	1069
Db	317	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	336
QY	1070	GCAGGAGTGACGAAGAGGAGAGGCGCAGAGCGATGGAGTACTCGAGTAGTGACAAC	1129
Db	337	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	356
QY	1130	AATGAAGAGGTAGTAGTCAAAAGTGTCAAAAGGAGCAGCTTGGAAGAACTTACTTAAGCAGCT	1189
Db	357	AsnGluGlyValIleValLysValSerLysGlnHisValGluGluLeuThrLysHisAla	376
QY	1190	AAATCCGCTCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCAACCAACCCAATCAACTTG	1249
Db	377	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	396
QY	1250	AGAGAAGCGAGCCGATCTTCTTAAACAATTTCGGGAAGTTATTGTAGGTGAAGCCAGAC	1309
Db	397	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	416
QY	1310	AAGAAGAACCCCGACGCTTCAGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
Db	417	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	436
QY	1370	GGAGCTTTGATGCTCCACACTTCAAATCAAGGCCATGTTATTCGTCTGCTCAACAAA	1429
Db	437	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	456
QY	1430	GGAACTGAAACCTTGAACCTCGTCGCTGAAGAAAGAGCAACAACAGAGGGGACGCGG	1489
Db	457	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgLysArg	476
QY	1490	GAAGAAGGAGGAGCAGAAGCAAGAAGAGAGGAGGAAGTAACACAGAGGTGCGTAGGTAC	1549
Db	477	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	496
QY	1550	ACAGCGAGGTTGAAGGAAGCGATGTTTCATCATATGCCAGCAGCTCATCAGTAGCCATC	1609
Db	497	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle	516
QY	1610	AACGCTTCTCCGAACCTCATCTCTGCTGGCTTCGGTATCAACGCTGAAAAACAACACAGA	1669
Db	517	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	536
QY	1670	ATCTTCTTCAGGTGATAAGGACAATGTGTATAGACCAGATAGAGAAGCAAGCGAAGGAT	1729
Db	537	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	556
QY	1730	TTAGCATCTCCGTGGGTGAACAAGTTGAGAGAGCTCATCAAAAACCAAGAGGATCT	1789
Db	557	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	576

Qy	1790	CAC TTGTGAGTCTCGTCTC CAATCTCAATCTCAATCTCCGTCGTCCTCTCTGAGAAGAG	1849
Dd	577	HIS PheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu	596
Qy	1850	TCTCTTGAGAAAGAGATCAAGAGGAGGAAAACAAGAGGGGAAGGTCCACCTCCTTTCA	1909
Dd	597	SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer	616
Qy	1910	ATT TTTGAAGCTTTTAAC	1927
Dd	617	IleLeuLysAlaPheAsn	622
RESULT 9			
US-09-731-221-78			
; Sequence 78, Application US/0973121			
; Patent No. US20020018778A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael			
; TITLE OF INVENTION: Passive Desensitization			
; FILE REFERENCE: 2002834-0103			
; CURRENT APPLICATION NUMBER: US/09/731,221			
; CURRENT FILING DATE: 2001-12-06			
; NUMBER OF SEQ ID NOS: 79			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 78			
; LENGTH: 634			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence:Arachis			
; OTHER INFORMATION: Hypogaea			
US-09-731-221-78			
Alignment Scores:			
Pred. No.:	4,47e-242	Length:	634
Score:	3052.00	Matches:	584
Percent Similarity:	96.86%	Conservative:	3
Best Local Similarity:	96.37%	Mismatches:	19
Query Match:	85.68%	Indels:	0
DB:	9	Gaps:	0
US-10-728-323-1 (1-2032) x US-09-731-221-78 (1-634)			
Qy	110	TCTGCAACGCATCCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTGCGCCCAG	169
Dd	16	SerSerThrHisAlaLysSerProTyrglnAlaLysThrGluAsnProCysAlaGln	35
Qy	170	AGGTGCTCTCAGAGTGTCTCAACAGGAACCGATGACTTGAAGCAAAGGCATCGCAGTCT	229
Dd	36	ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer	55
Qy	230	CGCTGACCAAGCTCGAGTATGATCTCTGTTGTGTCTATGATCCTCAGACACACTGGC	289
Dd	56	ArgCysThrIlysLeuGluTyraSpProArgCysAlaTyraSpProArgGlyHisThrGly	75
Qy	290	ACCACCAACCAACGTTCCCTCCAGGGGAGCGACACGTGGCCGCCAACCCCGAGACTAC	349
Dd	76	ThrThrAsnGlnArgSerProProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr	95
Qy	350	GATGATGACCGCCGTCACACCCCGAAGAGAGGAGGCGCGATGGGGACCACTGGACCG	409
Dd	96	AspAspAlaArgArgGlnProArgAlaGluGluGlyArgTrpGlyProAlaGlyPro	115
Qy	410	AGGAGCGTGAAGAGAGAAGNACTGGAGACACCAAGAGAGATTTGGAGGCGACCAAGT	469
Dd	116	ArgGluArgGluArgGluGluAspAlaArgGlnProArgGluAspTrpAlaArgProSer	135
Qy	470	CATCAGCAGCCACGGAAAAATAAGGCCCGAAGGAAGAGAGAACCAAGAGTGGGGAACA	529
Dd	136	HisGlnGlnProArgLysAlaArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	155
Qy	530	CCAGGTAGCCANTGTAGGGAAGAAACAATCTCGGAAACAACCCCTTTCTACTTCCTCAAGG	589

156 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 175
QY 590 CGGTTTAGACCCGCTACGGGAACCAAAACGGTAGATCCGGTCTCGAGAGTTTGAC 649
Db 176 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGluArgPheAsp 195
QY 650 CAAAGGTCAAGGAGTTTCAGAAATCCCAAGATCACCGTATTGTGCAGATCGAGGCCAA 709
Db 196 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 215
QY 710 CCTAACACTCTGTTCTTCCCAAGCAGCTGATGCTGATAAATCTTGTATTCCAGCAA 769
Db 216 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 235
QY 770 GGGCAAGCCACCGTAGCCGTAGCAATATGCAATAACAGAAAGCTTTTAATCTTGACGAG 829
Db 236 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 255
QY 830 GGCATGCACTCAGATCCCATCCGCTTTCATTTCTTACATCTTGAACCGCATGACAC 889
Db 256 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 275
QY 890 CAGAACTCTCAGAGTAGTAAATCTCCATGCCCGCTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 276 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnMetGluAsp 295
QY 950 TTCTTCCCGGAGCAGCCGAGACCAATATCTTCTTACGAGGCTTCAGCAGGAATACG 1009
Db 296 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheAlaArgAsnThr 315
QY 1010 TTGAGGCGCCTTCAATCCGGAATTCATGATACGAGGAGGTGCTGTAGAGGAAT 1069
Db 316 LeuGluAlaAlaPheAsnAlaGluAlaAsnGluIleArgArgValLeuLeuGluGluAsn 335
QY 1070 GCAGGAGGTGAGCAAGAGCAGAGCGCAGATGAGTACTCGGAGTAGTGAGAAC 1129
Db 336 AlaGlyGlyGluGlnGluAlaArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 355
QY 1130 AATGAAGAGTAGATGATCAAGTGTCAAGAGGACACGTTTGAAGAACTTACTAAGCAGCT 1189
Db 356 AsnGluGlyValIleValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 375
QY 1190 AATCCGCTCAAGAAAGCGCTCCGAGAGAGGGAGATATCACCAACCAATCACTTG 1249
Db 376 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProAlaAsnLeu 395
QY 1250 AGAAGAGCGAGCCCATCTTTCTAACTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 396 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuAlaGluValLysProAsp 415
QY 1310 AAGAAGACCCCGAGTTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 416 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 435
QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGGTTATCGTCTGTCACAA 1429
Db 436 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 455
QY 1430 GGAACCTGGAACCTTCAACTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGCGCGG 1489
Db 456 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnArgGlyArgArg 475
QY 1490 GAAGAAAGAGGAGCAAGAACGAGAGAGGGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
Db 476 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgAlaTyr 495
QY 1550 ACAGCAGGTTGAAGAAGCGGATGTTCATCATCGCAGCAGCTCATCCAGTAGCCATC 1609
Db 496 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 515
QY 1610 AACGCTTCTCGAACTCCATCTGCTTGGTTCGGTATCAACGCTGAAACCAACACACAGA 1669
Db 516 AsnAlaSerSerGluLeuAlaLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 535

1670 ATCTTCTCTTGCAGGTGATTAAGCAATGTGATAGACCAGATAGAGAAGCAAGCGAAGGAT 1729
Db 536 IlePheLeuAlaGlyAspAlaAspAsnValIleAspGlnIleGluLysGlnAlaLysAla 555
QY 1730 TTAGCATTTCCCTGGGTCCGGTCAACAAGTTGAGAAGCTCATCAAAAACAGAAAGGAATCT 1789
Db 556 LeuAlaAlaProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 575
QY 1790 CACTTTGTGAGTGCCTGCTCTCAATCTCAATCTCCGCTCGCTCTCTCAGAAAGAG 1849
Db 576 HisPheValAlaAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 595
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAAAAACCAAGAGGAGGAAGGTCCTCCTCTTCA 1909
Db 596 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 615
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 616 IleLeuLysAlaPheAsn 621
RESULT 10
US-09-331-631A-21
; Sequence 21, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULLIN23.001APC
; CURRENT APPLICATION NUMBER: US/09/331,631A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
US-09-331-631A-21
Alignment Scores:
Pred. No.: 3,55e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.53% Mismatches: 7
Query Match: 85.37% Indels: 18
DB: 9 Gaps: 7
US-10-728-323-1 (1-2032) x US-09-331-631A-21 (1-614)
QY 50 ATGAGAGGGAGGGTTTCTCCACTGATGCTGTGTAGGATCCTTGTCTGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGACGCATGCCAAGTCATCACCCTTACCGAAGAAAAACAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTCCCTCCAGAGTTGTCAACAGCAACCGGATGACTTGAACAAAAGCATGCCAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTGGCGCCCAACCCCGAGACTAC 349

```
Db      75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr  94
QY      350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGGATGGGGACCGACGCTGGACCG  409
Db      95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGluPro  114
QY      410 AGGGAGCGTGAAGAAGAAGACTGGAGACAACCAAGAGAAGATGGAGGGACCAAGT  469
Db      115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer  134
QY      470 CATCAGCACCGCGAATAATAGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  529
Db      135 HisGlnGlnProArgGlySerIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr  154
QY      530 CCAAGGTAGCGATGTAGGGAAGAAACATCTCCGAAACAACCCCTTCTACTTCCCTCAAGG  589
Db      155 ProGlySerGluValArgGluThrSerArgAsnAsnProPheTrpPheProSerArg  174
QY      590 CGGTTTAGCACCGCTACGGGAACCAAAACGGGTAGGATCCGGGTCTCGCAGAGGTTTGAC  649
Db      175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp  194
QY      650 CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAA  709
Db      195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg  214
QY      710 CTTAAACACTCTTGTCTCCCAAGCAGCGTGTGCTGATAAATCTCTGTTATTCAGACAA  769
Db      215 ProAsnThrLeuValLeuProLysHisIleAspAlaAspAsnIleLeuValIleGlnGln  234
QY      770 GGGCAAGCCACCGTGACCGTAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT  829
Db      235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu  254
QY      830 GGCATGCACTCAGAAATCCATCCGTTTCATCTTCTACATCTTGAAACCGCCATCAGAAC  889
Db      255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn  274
QY      890 CAGAACTCAGAGTAGCTAAATCTCATGCGCGTTAAACACACCCGCGAGTTTCAGGAT  949
Db      275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp  294
QY      950 TTTCTTCCCGGAGCAGCGGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAAATAGC  1009
Db      295 PhePheProAlaSerSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr  314
QY      1010 TTGGAGCGCGCTTCATCCGGAATTCATGAGATACGAGGGTGTGTTAGAAAGAGAAAT  1069
Db      315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn  334
QY      1070 GCAGGAGGTGACAGAGAGAGAGGAGCGGATGGAGTACTCGGAGTAGTGAGAAC  1129
Db      335 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp---  353
QY      1130 AATGAAGGAGTAGTAGTCAAAGTGTCAAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT  1189
Db      354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla  373
QY      1190 AATCCGCTCAAGAAAGCGTCCGAGAGAGGAGATATCACCAACCAATCAACTTG  1249
Db      374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu  392
QY      1250 AGAGAAGGCGAGCGCGATCTTTCTAACACTTTTGGAGAGTTATTGAGGTGAAGCCAGAC  1309
Db      393 ArgAspGlyGluProAspLysSerAsnAsnPheGlyArgLeuPheGluValLysProAsp  412
QY      1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTCTGTGTAGAGATCAAGAA  1369
Db      413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu  432
QY      1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGCATGGTTATCGTGTGCTCAACAA  1429
```

```
Db      433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValVal  452
QY      1430 CGAACTCGAAACCTTGAACCTCGTGGCTGTAGAAAGAGCAACACAGAGGGGACGCGG  1489
Db      453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg  472
QY      1490 GAA-----GAAGAGGAGGACGAGCAAGCAAGAGAGGAGGAGGAGTAACAGAGAGGTG  1540
Db      473 GluGlnGluTrpGluGluGluGluAspGluGluGluGluGluGluGluGluVal  492
QY      1541 CGTAGGTACACAGCGAGGTTGAAGCAAGCGATGTGTTTCATCATGCCAGCAGCTCATCCA  1600
Db      493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro  512
QY      1601 GTAGGCATCAACGCTTCCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTCAAAAC  1660
Db      513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn  532
QY      1661 AACCAAGAAATCTTCTTGCAGGTGATAGGACAATGTGATAGACAGATAGAGAACCA  1720
Db      533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln  552
QY      1721 CGCAAGGATTTAGCATTCCTCGGTCGGTGAACCAAGTTGAGAAGCTCATCAAAACCAAG  1780
Db      553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln  572
QY      1781 AAGGAATCTCACTTTGTGAGTGTGCTCTCAATCTCAATCTCAATCTCCGCTCGTCTCTCT  1840
Db      573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer-----  588
QY      1841 GAGAAAGATCTCTGAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG  1900
Db      589 -----SerProGluLysGluAspGlnGluGluGluGluGluGluGlyGlyLysGlyPro  605
QY      1901 CTCCTTTCAATTTTCAAGCGCTTTTAAAC  1927
Db      606 LeuLeuSerIleLeuLysAlaPheAsn  614

RESULT 11
US-09-847-208-27
; Sequence 27, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: UC67.002A
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-27

Alignment Scores:
Pred. No.:      3.55e-241      Length:      614
Score:          3041.00      Matches:      595
Percent Similarity: 96.03%      Conservative: 9
Best Local Similarity: 94.59%      Mismatches: 7
Query Match:      85.37%      Indels:      18
Db:              10          Gaps:      7

US-10-728-323-1 (1-2032) x US-09-847-208-27 (1-614)
QY      50 ATGAGAGGAGGGTTTCTCCACTGATGTGTGTGTAGGATCCTTGTCTCGCTTCAGTT  109
Db      1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValValValValVal  20
```

110 TCTGCAACGATCCCAAGTCATCACCTTACCAGAAACAGAGAACCCCTGCGCCAG 169
Db SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
170 AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT 229
Db ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
230 CGTGTGACCAAGTCGAGTATGATCTCGTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGCGCCCAACCCGGAGACTAC 349
Db AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
350 GATGATGACCGCGTCAACCCCAAGAGAGAGGAGCGGATGGGGACACGATGGACCG 409
Db AspAspAspArgArgGlnProArgArgGluGlyArgTrpGlyProAlaGluPro 114
410 AGGGAGGTGAAGAGAGAGACTCGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
Db ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 134
470 CATCAGCAGCCAGGAAATAGGCCCGGAGGAGAGAGAGAGAAACAAGAGTGGGGAACA 529
Db HisGlnGlnProArgLysIleArgProGluGlyArgGluGlnGlnTrpGlyThr 154
530 CAGGTAGCCATGTAGGGAAGAAACATCTCGAAACAACCTTCTACTTCCCGTCAAGG 589
Db ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
590 CGGTTTAGCACCGCTACGGGAACCAAAACCGTAGGATCGGGTCTCGAGAGTTTGAC 649
Db ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
650 CAAAGTCAAGCGAGTTTCAGATCTCCAGATCCAGATCCAGTATCTGTCAGATCGAGCCAA 709
Db GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
710 CTTAACACTCTTGTCTTCCAGACACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
Db ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
770 GGGCAAGCCACCGTGACCGTAGCAATGCAATTAACAGAAAGAGCTTTAATCTTGACGAG 829
Db GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
830 GGCCTAGCATCAGATCCCATCCGTTTCATTTCTCTACATCTTGAAACCGCCATGACAAC 889
Db GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
890 CAGAACTCAGAGTAGTAAATCTCCATGCGCCGTTAAACACACCCCGCCAGTTTGAGGAT 949
Db GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTACTTGAGGGCTTCAGCAGGAATACG 1009
Db PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGGTGTCTGTAGAAGAGAT 1069
Db LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
1070 GCAGGAGGTGAGCAAGAGGAGAGAGCGGATGAGTACTCGGAGTAGTGAGAAC 1129
Db AlaGlyGlyGlnGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
1130 AATGAAGGATGATGATCAAGTGTCAAAGGAGCACGTTGAAGACTTACTTAAGCACGCT 1189
Db AsnGluGlyValIleValLysValSerLysGluHisValGlnGlnLeuThrLysHisAla 373
1190 AATCCGCTCTCAAGAAAGGCTCCGAGAGAGGAGATATACCAACCCCAATCAACTTG 1249

374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAAGCGAGCCCGATCTTTTAACTTTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAAGCCCGAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGATCAAGAAG 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGTCGTCAACAAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAATCGGAACCTTGAACCTCGTGTGTAAGAAAGAGCAACAACAGAGGGGACCGCG 1489
Db 453 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGGAGCAGACGAGAGAGAGAGAGGAGTAAACAGAGAGGTG 1540
Db 473 GluGlnGluTrpGluGluGluAspGluGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCAGGAGTTGAAGGAGCGGATGTTTCATCATGCCAGCAGCTCATCCA 1600
Db 493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCCGAATCCATCTGCTTGGTTCGGTATCAACGCTGAAAAC 1660
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAAGATCTTCTTCAGGTATTAAGGACATGTGATAGACAGATAGAGAGCAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTGCAAGTTGAGAAGCTCATCAAAACACAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCA 1900
Db 589 -----SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTTCAATTTTGAAGGCTTTTAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614
RESULT 12
US-10-147-095-21
; Sequence 21, Application US/10147095
; Publication No. US20030171274A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULN23.001APC
; CURRENT APPLICATION NUMBER: US/10/147,095
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21

; LENGTH: 614									
; TYPE: PRT									
; ORGANISM: Peanut									
US-10-147-095-21									
Alignment Scores:									
Pred. No.:	3.55e-241	Length:	614						
Score:	3041.00	Matches:	595						
Percent Similarity:	96.03%	Conservative:	9						
Best Local Similarity:	94.59%	Mismatches:	7						
Query Match:	85.37%	Indels:	18						
DB:	14	Gaps:	7						
US-10-728-323-1 (1-2032) x US-10-147-095-21 (1-614)									
QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTCTAGGATCCTTGTCTCGGCTTCAGTT	109						
DB	1	MeArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal	20						
QY	110	TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAGAAACAGAGAACCCCTGCGCCAG	169						
DB	21	SerAlaThrGlnAlaLys--SerProTyr---ArgLysThrGluAsnProCysAlaGln	38						
QY	170	AGTGCCCTCAGAGTTGTCAACAGGAACGGATGACTTGAAGCAAAAGGCATCGAGTCT	229						
DB	39	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	58						
QY	230	CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC	289						
DB	59	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly	74						
QY	290	ACCACCAACCAAGTTCCCTCCAGGGAGCGGACACGTTGGCGGCCAACCCGGAGACTAC	349						
DB	75	AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	94						
QY	350	GATGATGACCGCGCTCAACCCGAGAGAGAGAGAGAGCGGATGGGACCACTGGACCG	409						
DB	95	AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGluPro	114						
QY	410	AGGGAGCGTGAAGAGAGAGCAAGACTGGAGACAAACCAAGAGAGATTCGAGGCGCAAGT	469						
DB	115	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	134						
QY	470	CATCAGCACCACGGAAATAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529						
DB	135	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	154						
QY	530	CGAGTAGCCATGTGAGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAGG	589						
DB	155	ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	174						
QY	590	CGGTTTTCAGCACCCTACGGGAACCAAAACGCTAGGATCCGGTCTCGCAGAGTTTCAC	649						
DB	175	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	194						
QY	650	CAAAGGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCCAA	709						
DB	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214						
QY	710	CTTAACACTCTTGTGTTCTCCCAAGCACGCTGATGCTGATAACATCTTGTGTTATCCAGCA	769						
DB	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234						
QY	770	GGGCAAGCCACCGTACCGTACCAATGGCAATACAGAAAGAGCTTTAATCTTCACCGAG	829						
DB	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254						
QY	830	GGCCATGCACTCAGAAATCCCATCCGGTTTTCATTTCTTACATCTTGAACCGCCATGCAAC	889						
DB	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274						
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCCGGTTTACACACCCGCCAGTTTGAGGAT	949						

RESULT 13

US-10-100-303A-8

; Sequence 8, Application US/10100303A

DB	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTCTTCCCGGAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTCAGCAGGAATACG	1009
DB	295	PhePheProAlaSerArgAspGlnSerSerYrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTGAGGCGCGCTTCAATGCCGAATTTCAATGAGATACGAGGGTCTGTGTAGAGAGAT	1069
DB	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	334
QY	1070	GCAGAGGTGAGCAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1129
DB	335	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
QY	1130	AATGAAGGAGTCATAGTCAAAAGTGTCAAAGGAGGAGGAGGAGGAGGAGGAGGAGG	1189
DB	354	AsnGluGlyValIleValLysValSerLysGluHisValGlnGlnLeuThrLysHisAla	373
QY	1190	AAATCCGTCTCAAAGAAAGGCTCCGAAAGAGAGGAGGAGATATCACCACCCCAATCA	1249
DB	374	LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu	392
QY	1250	AGAGAAGCGGAGCCCGATCTTTTAAACAACTTTGGGAAGTTATTGAGGTGAAGCCAG	1309
DB	393	ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp	412
QY	1310	AAGAAGAACCCCGCTTCCAGGACCTGACATGATGCTCCTGTGTAGAGATCAAGAA	1369
DB	413	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLysGlu	432
QY	1370	GGAGCTTTGATGCTCCCACTTCAAACCTCAAAGGCCATGGTTTATCGTCTGTCACAAA	1429
DB	433	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	452
QY	1430	GGAATCGAAACCTTGAACTCGTGTCTGTAGAAAGAGCAACACAGAGGAGCGGG	1489
DB	453	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	472
QY	1490	GAA-----GAAGAGGAGGAGCAAGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG	1540
DB	473	GluGlnGluTrpGluGluGluGluGluGluGluGluGluGluGluGluGluGluVal	492
QY	1541	CGTAGGTACACAGCGAGGTTCAAGGAAGCGATGTTCATCATGCTCCAGCAGCTCATCA	1600
DB	493	ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro	512
QY	1601	GTAGCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAC	1660
DB	513	ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn	532
QY	1661	AACACAGAACTTCTTCCAGGTGATAAAGGACAAATGTGTAGACACGATAGAGAGCAA	1720
DB	533	AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln	552
QY	1721	CGCAAGGATTTAGCATTCCTCGGGTCCGGGTGAACAAGTTGAGAAGCTCATCAAAAC	1780
DB	553	AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln	572
QY	1781	AAGGAATCTCACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCTCTCT	1840
DB	573	ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer-----	588
QY	1841	GAGAAAGAGTCTCTGAGAAAGAGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGG	1900
DB	589	-----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyPro	605
QY	1901	CTCCTTTTCAATTTGAAGCTTTTAAAC	1927
DB	606	LeuLeuSerIleLeuLysAlaPheAsn	614

Publication No. US20030202980A1
GENERAL INFORMATION:
APPLICANT: Caplan, et al.
TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
FILE REFERENCE: 2002834-0166
CURRENT APPLICATION NUMBER: US/10/100,303A
CURRENT FILING DATE: 2002-03-18
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 614
TYPE: PRT
ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-8

Alignment Scores:
Pred. No.: 3-55e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.03% Conservative: 9
Best Local Similarity: 94.59% Mismatches: 7
Query Match: 85.37% Indels: 18
DB: 15 Gaps: 7

US-10-728-323-1 (1-2032) x US-10-100-303A-8 (1-614)

QY	50	ATGAGAGGGAGGTTTCTCCACGTAGTGTCTTCTAGGATCTTGTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCGCAACGCATCCCAAGTCATCACCTTACCAGAGAAACACAGAACCCCTCGGCCAG	169
DB	21	SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGlnAsnProCysAlaGln	38
QY	170	AGTGTGCTCCAGAGTGTCTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCCGAGTCT	229
DB	39	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	58
QY	230	CGTGTCAACAAAGTCAGATATGATCTCTGTGTGTCTATGATCTCTCGAGACACATGCG	289
DB	59	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly	74
QY	290	ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGGAGACTAC	349
DB	75	AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	94
QY	350	GATGATGACCGCGCTCAACCCCAAGAGAGGAGGCGCGATGGGGACCGCTGGACCG	409
DB	95	AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro	114
QY	410	AGGAGCGTGAAGAAGAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT	469
DB	115	ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer	134
QY	470	CATCAGCCCGGAAATAGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	135	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr	154
QY	530	CAAGGTAGCCATGTAGAGGAAGAAACATCTCGGAACAAACCTTCTACTTCCTCCGTC	589
DB	155	ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	174
QY	590	CGGTTTAGCACCGCTACCGGAAACCAAAACCGTGGATCCGGTCTCGAGAGGTTTGAC	649
DB	175	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	194
QY	650	CAAGGTCAAGGAGTTCAGATCTCCAGATCACCGTATCTGTGAGATCGAGGCCAAA	709
DB	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
QY	710	CCTAACACTCTTGTCTCCAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	769
DB	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
QY	770	GGGCAAGCCACCGTGTAGCCCTAGCAATGCAATGCAATGCAATGCAATGCAATGCAAT	829
DB	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
QY	830	GGCCATGCACTCAGAAATCCCATCCGTTTTCATTTCTTCTACATCTTGAACCGCATGACAAC	889
DB	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
QY	890	CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAAGAT	949
DB	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTTCTCCCGCGAGAGCCGAGACCAATCATCTTCTTCTGAGGGCTTCAGCAGGAATACG	1009
DB	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTTGGAGCCGCTTCAATGCGGAATTCATGATAGATACGAGGGTGTCTTGAAGAGAAAT	1069
DB	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	334
QY	1070	GCAGGAGGTGAGCAAG	1129
DB	335	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
QY	1130	AATGAGGAGTGTATGTCAAAGTGTCAAAGAGACGCTTTGAAGAACTTATTAAGCAAGCT	1189
DB	354	AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla	373
QY	1190	AAATCGCTCTCAAGAAAGGCTCCCAAGAGAGAGAGAGAGATATCACCAACCAATCACTTG	1249
DB	374	LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu	392
QY	1250	AGAGAGGCGAGCCCGATCTTTCTAACTTTGGGAGCTTATTTGAGGTGAAGCCAGAC	1309
DB	393	ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp	412
QY	1310	AAGAAGAACCCCGCTCAGGACCTCGACATGATGCTCCTCTGTGTAGAGATCAAAGAA	1369
DB	413	LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu	432
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCGATGGTTATCGTCTCGTCAACAA	1429
DB	433	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	452
QY	1430	GGAACTGGAAACCTTCAACTCGTGTCTGTAAAGAAAGAGCAACACAGAGGGAGCGCG	1489
DB	453	GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	472
QY	1490	GAA-----GAG	1540
DB	473	GluGlnGluTyrGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	492
QY	1541	CGTAGGTACAGAGGTTGAGAGAGCGGATGTTTCATCATCGTCCAGCAGCTCATCCA	1600
DB	493	ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro	512
QY	1601	GTAGCCATCAACCGCTTCTCCGAACTCCATCTGTGCTGGCTTCGGTATCAACGCTGAAAAC	1660
DB	513	ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn	532
QY	1661	AACCAAGAAATCTTCTTCTGAGGTGATAGGCAATGTGATAGACAGATAGAGAGCAAA	1720
DB	533	AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln	552
QY	1721	CGAAGGATTTAGCATTCCTCGGTCGGGTGAACAGTTGAGAAGCTCATCAAAACCCAG	1780
DB	553	AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln	572
QY	1781	AAGGAATCTCACTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCAATCTCAATCT	1840
DB	573	ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer-----	588

QY 1841 GAGAAAGAGTCTCTCGAAGAGAGGATCAAGAGGAGGAAACCAAGAGGGAGGGTCCA 1900
Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605

QY 1901 CTCCTTTCAATTTTGAAGGCTTTTAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 14
US-10-100-303A-55
; Sequence 55, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-55

Alignment Scores:
Pred. No.: 1,04e-105 Length: 268
Score: 1394.00 Matches: 268
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.14% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-100-303A-55 (1-268)

QY 485 AAAATAAGCCCGGAGGAAGAGAGAGAGTGGGGGACACACAGGTAGCCATGTG 544
Db 1 LysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThrProGlySerHisVal 20

QY 545 AGGGAAGAAACATCTCGGAACACCCCTTCTACTCCCTCAAGGCGGTTTAGCCCGC 604
Db 21 ArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArgArgPheSerThrArg 40

QY 605 TACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTGACCAAGGTCAGGCGAG 664
Db 41 TyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAspGlnArgSerArgGln 60

QY 665 TTTTCAGAAATCTCCAGAAATCACCGGTATGTGCAGATCGAGGCCAAACCTAACACTCTTGT 724
Db 61 PheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLysProAsnThrLeuVal 80

QY 725 CTTCCCAAGCAGCTGATGCTGATACATCTTGTATCCAGAGGGCAAGCCACCGTG 784
Db 81 LeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGlnGlyGlnAlaThrVal 100

QY 785 ACCGTAGCAAAATGGCAATAACAGAAAGAGCTTAACTTTGACGAGGGCCATGCTCAGA 844
Db 101 ThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGluGlyHisAlaLeuArg 120

QY 845 ATCCCATCCGGTTTCATTTTCTACATCTTTGAACCGCATGACAAACAGAAACCTCAGAGTA 904
Db 121 IleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsnGlnAsnLeuArgVal 140

QY 905 GCTAAATCTCCATGCCCTTAAACACACCGGCGAGTTTGAGATTTCTCCCGGGAGC 964
Db 141 AlaLysIleSerMetProValAsnThrProGlyGlnPheGluAspPhePheProAlaSer 160

QY 965 AGCCGAGACCAATCATCTTACTTTGAGGGCTTCAGCAGGAATACGTTGAGGCGCCCTTC 1024
Db 161 SerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThrLeuGluAlaPhe 180

QY 1025 AATGCGGAATTAATAGATACGGAGGTGCTGTTTAGAAGAGAAATGCAGGAGGTGAGCAA 1084
Db 181 AenAlaGluPheAsnGluIleArgArgValLeuLeuGluAsnAlaGlyGlyGluGln 200

QY 1085 AAGGAGAGAGGCGAGGCGATGGAGTACTGGAGTAGTGAGAACAAATGAAGGAGTGATA 1144
Db 201 GluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIle 220

QY 1145 GTCRAAGTGTCAAAGGAGCACGTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTCCAAG 1204
Db 221 VallysValSerLysGluHisValGluLeuLeuThrLysHisAlaLysSerValSerLys 240

QY 1205 AAAGGCTCCGAAGAGAGGAGATATCACCAACCCAATCAACTTGAGAGAGCGGAGCCC 1264
Db 241 LysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeuArgGluGlyGluPro 260

QY 1265 GATCTTTCTAACAACTTTGGGAAG 1288
Db 261 AspLeuSerAsnAsnPheGlyLys 268

RESULT 15
US-10-424-599-153195
; Sequence 153195, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 153195
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_10935C.1.pep
US-10-424-599-153195

Alignment Scores:
Pred. No.: 2.43e-96 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.37% Conservative: 108
Best Local Similarity: 43.05% Mismatches: 173
Query Match: 35.99% Indels: 96
DB: 15 Gaps: 18

US-10-728-323-1 (1-2032) x US-10-424-599-153195 (1-605)

QY 50 ATGAGAGGAGGAGTTTCTCCACTGATGCTGTTGTAGGATCCTTGTCTGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 19

QY 110 TCTGCAACGCATGCCCAAGTCATCACCTTACCAGAAAGAAACAGAGAACCCCTGCGCCAG 169
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36

QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGCACTTTGAAGCAAAAGCATGCGAGTCT 229
Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56

QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGlyGluIleProArg--- 75

QY 263 GTCTATGATCTCGAGGACACACTGGCACCAACCAACGTTCCCTCCAGGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84

```
QY 323 ACACGTGGCCGCGAACCAGGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnProGlyGluLysGluAspGluGlnProArgProIle 104
QY 359 -----GCCGCTCAACCCCGAAGAGAG-----GAAGGAGGCCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGlnArgGluGlnGlu 124
QY 392 TGG-----GGACCAGCTGGACCGGAGGAGCGGTGAAGAGAAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
QY 431 GACTTGGAGACAACCAAGAGAAGAT-----TGGAGGCGACCAAGTCATCAGCAGCCACGG 484
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
QY 485 AAAATAAGGCCGAAGAGAGAGAAACAAGATGGGGAAACACACGAGTAGCCATGTG 544
Db 165 ArgLysGlnGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluAsp 184
QY 545 AGGGA-----GAAACATCTCGAACAACCCCTTCTACTCCGCTCAAGGCGGTTAGC 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCCGCTACGGGAACCAAAACGGTAGATCCGGCTCTCGAGAGTTTGACCAAGGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCCAGGCCAAACCTAACACT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTTGTTCTTCCCAAGCAGCTGATGTGATAACATCTTGTTATCCAGCAAGGCGCAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTGACCGTAGCAATGCCAATAACAGAAAGAGCTTTAATCTTCAGGAGGCCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGAAATCCCATCCGCTTTCATTCTCATCTCTGAACCGCCATGACCAACAGAACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGCTAAATCTCCATGCCGCTTAACACACCGCGCAGTTTGAGATTTCTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
QY 959 GCGAGCAGCGGACCAATCATCTACTTCAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 325 SerThrThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAATTCAATGAGATACGAGGGTCTGTTAGAGAAGATCGCAGGAGT 1078
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GACCAAGAGAGAGAGCGGACGCGATGAGTACTCGAGTAGTCAGAACATGAAGGA 1138
Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAGAGGACGAGTTGAAGAACTTACTAAGCACCGCTAAATCCGTC 1198
Db 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY 1199 TCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCATCACTTGTAGAGAAGGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGACAGAAGAAC 1318
Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCCAGCTTCAGGACCTGGACATGCTCACCTCTGTGTAGATCAAAAGAGGAGCTTTG 1378
```

```
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAGGCGCATGTTATCGTCGTCAACAAAGGAAGACTGGA 1438
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAATCTCGTGGCTCTAAGAAAGAGCAACAACAGAGGGGACGCGGGGAAGAG 1498
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAAGACGAAGAGGAGGGAAGTAACAGAGAGTGCCTAGGTACACAGCGAGG 1558
Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAAGGCGATGTGTTCATCATGCGCAGAGCTCATCCAGTAGCCATCAACGCTTCC 1618
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValAlaAsnAlaThr 519
QY 1619 TCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGAAATCTTCCCT 1678
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGCAATCTGTATAGCACCATAGAGACGACGACGAGGATTTAGCATTC 1738
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CCTGGGTCGGGTGAACAAGTTGAGAAGCTCATCAAAAAACAGAAAGAAATCTCACTTTGTG 1798
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCCGTCGTCTCTCGAGAAAGAGTCTCTCTGAG 1858
Db 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGAGGAGAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAG 1918
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
Db 603 AlaPhe 604
```

Search completed: August 24, 2005, 10:52:48

Job time : 167.012 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 2122.81 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcacatactagtagcccc.....taaaaagatcatgtttttgtt 717

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	717	100.0	717	6	AR257468 Sequence
2	717	100.0	717	8	L77197 Arachis hyp
3	713.8	99.6	743	6	AR257462 Sequence
4	705	98.3	810	8	AY722689 Arachis h
5	644.6	89.9	842	8	AY581853 Arachis h
6	598.8	83.5	678	8	AY581854 Arachis h
7	474	66.1	474	6	BD107899 Methods a
8	472.4	65.9	474	6	AX155333 Sequence
9	470.8	65.7	474	6	BD172108 Peanut al
10	462	64.4	472	8	AY007229 Arachis h
11	462	64.4	682	6	AX148740 Sequence
12	462	64.4	1162	6	AX148738 Sequence
13	419.4	58.5	519	8	AY158467 Arachis h
14	415.2	57.9	510	8	AY117434 Arachis h
15	404.8	56.5	502	8	AF366560 Arachis h
16	220.8	30.8	634	8	AF366561 Arachis h
17	202.2	28.2	742	8	AY722690 Arachis h
18	197.8	27.6	627	8	AF092846 Arachis h
19	114.8	16.0	712	8	AF091737 Arachis h

20	108	15.1	581	8	AY722691 Arachis h
21	80	11.2	80	6	AX148744 Sequence
22	73.2	10.2	684	8	LACONGLD X53523 L. angustif
23	65.6	9.1	777	6	AR066463 Sequence
24	65.6	9.1	777	8	GMU71195 Glycine max
25	64	8.9	777	6	AR066464 Sequence
26	63.4	8.8	770	6	AR107077 Sequence
27	63.4	8.8	770	6	AR304633 Sequence
28	63.4	8.8	770	6	BD075791 Lunacin p
29	63.4	8.8	770	6	AF005030 Glycine m
30	62	8.6	62	6	AX148745 Sequence
31	59.2	8.3	477	6	AX839974 Sequence
32	59.2	8.3	477	6	AX840499 Sequence
33	58.8	8.2	7218	6	I66494 Sequence 14
34	57.6	8.0	82139	3	AC115684 Dictyoste
35	56.4	7.9	127640	8	AC146555 Medicago
36	55	7.7	160542	2	CR627487 Danio rer
37	54.4	7.6	468	6	AX839972 Sequence
38	54.4	7.6	468	6	AX840497 Sequence
39	54.4	7.6	723	6	AR066462 Sequence
40	54.4	7.6	724	8	GMU71194 U71194 Glycine max
41	53.6	7.5	174311	2	CR678065 Danio rer
42	52.8	7.4	168571	9	CNS01DUF AL133249 BAC sequ
43	52.8	7.4	197146	5	BX470199 Zebrafish
44	52.8	7.4	230976	2	BX927201 Danio rer
45	52.6	7.3	199551	2	AC006281 Plasmodiu

ALIGNMENTS

RESULT 1
AR257468
LOCUS AR257468 717 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 20 from patent US 6486311.
ACCESSION AR257468
VERSION AR257468.1 GI:27307481
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 717)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helim,R.M. and Bannon,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 20 26-NOV-2002;
FEATURES
Location/Qualifiers
source 1..717
/organism="unknown"
/mol_type="genomic DNA"

Query Match	Score	DB	Length
Best Local Similarity	100.0%	717;	DB 6; Length 717;
Matches	717;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	GCTCACCATCTAGTAGTACCTCGCCCTTTCTCTCGTGCCTCCACGCAATCTCGGAGGCA	60
Db	1	GCTCACCATCTAGTAGTACCTCGCCCTTTCTCTCGTGCCTCCACGCAATCTCGGAGGCA	60
Qy	61	GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAAGCTGAG	120
Db	61	GCAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGAAGCTGAG	120
Qy	121	GCCTCGGAGCAACATCTCATGCAGAGAGATCCAACTGACGAGATTCATATCAACGGGA	180
Db	121	GCCTCGGAGCAACATCTCATGCAGAGAGATCCAACTGACGAGATTCATATCAACGGGA	180
Qy	181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGG	240
Db	181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGG	240
Qy	241	ATCCTCTAGCACCAGAGAGAGGTGTGCAATGAGCTGAACAGGATTTTGAGAACCAACCAAG	300

```
Db 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGTGAACGAGTTTGAGAAACAACCAAG 300
QY 301 GTGCATGTGCGAGGCAATGCAACAGATCATGAGAAACCGAGAGATAGGTTGCGAGGGAG 360
Db 301 GTGCATGTGCGAGGCAATGCAACAGATCATGAGAAACCGAGAGATAGGTTGCGAGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCCTTAG 420
QY 421 GGCAACCAAGCGTTGCGAATTGGAACGTGCAAGTGGCGGAGAGACAGATCTAAACACC 480
Db 421 GGCAACCAAGCGTTGCGAATTGGAACGTGCAAGTGGCGGAGAGACAGATCTAAACACC 480
QY 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
QY 541 GTTATGTTTATGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
Db 541 GTTATGTTTATGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGCGGCTGTCTATGT 660
Db 601 ACTAAGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGCGGCTGTCTATGT 660
QY 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAAAGAAAGAAAGAAAGAAAGAAAGAAAG 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAAAGAAAGAAAGAAAGAAAGAAAGAAAG 717

RESULT 2
ARQALLII
LOCUS Arachis hypogaea (clone Ara h II p38) allergen II gene, polyA 28-MAR-1996
DEFINITION signal.
VERSION L77197
KEYWORDS allergen; conglutinin; seed storage protein.
SOURCE Arachis hypogaea
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 717)
AUTHORS Stanley,J.S.
TITLE The major peanut allergen Ara h II is a seed storage protein with
JOURNAL multiple IgE-binding epitopes
COMMENT Unpublished (1996)
ORIGINAL SOURCE text: Arachis hypogaea (strain Florunner) (clone:
FEATURES
source Location/Qualifiers
polyA_signal 562..567
ORIGIN
Query Match 100.0%; Score 717; DB 8; Length 717;
Best Local Similarity 100.0%; Pred.No.1.7e-171; Indels 0; Gaps 0;
Matches 717; Conservative 0; Mismatches 0;

QY 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGATCTGCGAGGCA 60
Db 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGATCTGCGAGGCA 60
QY 61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120
```

```
Db 61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120
QY 121 GCCTTCGCGAGCAACATCTCATGCAGAAGATCAAAGTGCAGGAGATTCAATATGAACGGGA 180
Db 121 GCCTTCGCGAGCAACATCTCATGCAGAAGATCAAAGTGCAGGAGATTCAATATGAACGGGA 180
QY 181 CCCGTACAGCCCTAGTACGATCCGCTACAGCCCTAGTCCATATGATCGGAGAGGGCGCTGG 240
Db 181 CCCGTACAGCCCTAGTACGATCCGCTACAGCCCTAGTCCATATGATCGGAGAGGGCGCTGG 240
QY 241 ATCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
Db 241 ATCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
QY 301 GTGCATGTGCGAGGCAATGCAACAGATCATGAGAAACCGAGAGATAGGTTGCGAGGGAG 360
Db 301 GTGCATGTGCGAGGCAATGCAACAGATCATGAGAAACCGAGAGATAGGTTGCGAGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTTGCCCTCAACAGTGGCGCCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTTGCCCTCAACAGTGGCGCCTTAG 420
QY 421 GGCACCAAGCGTTGCGAATTGGAACGTGCAAGTGGCGGAGAGACAGATCTAAACACC 480
Db 421 GGCACCAAGCGTTGCGAATTGGAACGTGCAAGTGGCGGAGAGACAGATCTAAACACC 480
QY 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
QY 541 GTTATGTTTATGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
Db 541 GTTATGTTTATGTTTGGTAAATAAAGATCATCACTATATGAATGTGTTGATCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGCGGCTGTCTATGT 660
Db 601 ACTAAGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGCGGCTGTCTATGT 660
QY 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAAAGAAAGAAAGAAAGAAAGAAAGAAAG 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAAAGAAAGAAAGAAAGAAAGAAAGAAAG 717

RESULT 3
AR257462
LOCUS Ar257462
DEFINITION Sequence 1 from patent US 6486311.
ACCESSION AR257462
VERSION AR257462.1
KEYWORDS GI:27307475
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 743)
AUTHORS Burks,A.W. Jr., Stanley,J.S., Cockrell,G., King,N.E., Sampson,H.A.,
Helm,R.M. and Bannon,G.A.
TITLE Peanut allergens and methods
JOURNAL Patent: US 6486311-A 1 26-NOV-2002;
FEATURES
source Location/Qualifiers
polyA_signal 1..743
ORIGIN /organism="unknown"
/mol_type="genomic DNA"

Query Match 99.6%; Score 713.8; DB 6; Length 743;
Best Local Similarity 99.7%; Pred.No.1.1e-170; Indels 0; Gaps 0;
Matches 715; Conservative 0; Mismatches 2;

QY 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGATCTGCGAGGCA 60
Db 6 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTGCCTCCACGATCTGCGAGGCA 65
QY 61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120
```

Db	66		GCAGTGGGAACTTCCAGGAGACAGAAGATGCCAGAGGCAGCTTCGAGAGGGCGGAACCTGAG	125
Qy	121		GCCCTGCGAGCAACATCTCTCATGTCAGAAGATCCAAAGTGCACGAGGATTCATATGAACGGGA	180
Db	126		GCCCTGCGAGCAACATCTCTCATGTCAGAAGATCCAAAGTGCACGAGGATTCATATGAACGGGA	185
Qy	181		CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGCGCTGG	240
Db	186		CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGCGCTGG	245
Qy	241		ATCCTCTCAGCACCAAGAGAGGTGTGCAATCAGCTGAACGAGTTTGAGAAACAACCAAG	300
Db	246		ATCCTCTCAGCACCAAGAGAGGTGTGCAATCAGCTGAACGAGTTTGAGAAACAACCAAG	305
Qy	301		GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACACAGACCGATAGGTTCCAGGGGAG	360
Db	306		GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACACAGACCGATAGGTTCCAGGGGAG	365
Qy	361		GCAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACTTGCCCTCAAACAGTGGCGCTTAG	420
Db	366		GCAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACTTGCCCTCAAACAGTGGCGCTTAG	425
Qy	421		GGCACACAGCGTTGCGACTTGGACGTGCGAAGTGGCGGCGAGAGACAGATACTAAACACC	480
Db	426		GGCACACAGCGTTGCGACTTGGACGTGCGAAGTGGCGGCGAGAGACAGATACTAAACACC	485
Qy	481		TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	540
Db	486		TATCTCAAAAAAGAAAAAGAAAAAGAAAAATAGCTTATATATAAGCTATTATCTATG	545
Qy	541		GTTATGTTTGTGTTGTPAATAATAAGATCATCACTATATGAATGTGTGATCGTGTA	600
Db	546		GTTATGTTTGTGTTGTPAATAATAAGATCATCACTATATGAATGTGTGATCGTGTA	605
Qy	601		ACTAAGCAAGCTTAGGTTATATAGACACCTTTAGAGTGCCTTTTAGCGTGTGCTATGT	660
Db	606		ACTAAGCAAGCTTAGGTTATATAGACACCTTTAGAGTGCCTTTTAGCGTGTGCTATGT	665
Qy	661		TTTGTTGCTGCAGAGTTGTAAACCATCTTTGAAATAATATAAAAGATCATGTTTTGTT	717
Db	666		TTTGTTGCTGCAGAGTTGTAAACCATCTTTGAAATAATATAAAAGATCATGTTTTGTT	722

RESULT 4	AY722689	810 bp	linear	PLN 19-SEP-2004
LOCUS	AY722689			
DEFINITION	Arachis hypogaea 2S protein 1 mRNA, partial cds.			
ACCESSION	AY722689			
VERSION	AY722689.1			
	GI:52001226			

[illegible]

ORIGIN

Query Match	98.3%	Score 705;	DB 8;	Length 810;
Best Local Similarity	99.9%;	Pred. No. 1.9e-168;		
Matches 716;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
Qy	1	GCTCACCATACTAGTAGCCCTCGCCCTTTTCCCTCTCGCTGCCACGCATCTCGAGGCA	60	
Db	68	GCTCACCATACTAGTAGCCCTCGCCCTTTTCCCTCTCGCTGCCACGCATCTCGAGGCA	127	
Qy	61	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAG	120	
Db	128	GCAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAG	187	
Qy	121	GCCTCTGGAGCAACATCTCATGCAAGAAGATCCAAAGTGCAGAGGATTCATATGAACGGGA	180	
Db	188	GCCTCTGGAGCAACATCTCATGCAAGAAGATCCAAAGTGCAGAGGATTCATATGAACGGGA	247	
Qy	181	CCGCTACAGCCCTAGTTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGAGCGCTGG	240	
Db	248	CCGCTACAGCCCTAGTTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGAGCGCTGG	307	
Qy	241	ATCCTCTCAGCAACAAGAGAGGTGTTGCAATGAGCTGAACAGAGTTTGAGAAACAACCAAG	300	
Db	308	ATCCTCTCAGCAACAAGAGAGGTGTTGCAATGAGCTGAACAGAGTTTGAGAAACAACCAAG	367	
Qy	301	GTGCATGTGGAGGCATTTGCAACAGATCATGAGAGAACAGAGGATAGTTTGCAGGGGAG	360	
Db	368	GTGCATGTGGAGGCATTTGCAACAGATCATGAGAGAACAGAGGATAGTTTGCAGGGGAG	427	
Qy	361	GCAACAGGAGCAACAGTTTCAAGAGGAGGCTCAGCACTTCGCTCAACAGTGCAGCCCTTAG	420	
Db	428	GCAACAGGAGCAACAGTTTCAAGAGGAGGCTCAGCACTTCGCTCAACAGTGCAGCCCTTAG	487	
Qy	421	GGCACCAACGCTTGGACGCTGGAAGCTGCGAAGTGGCGGAGACAGATACATAAACACC	480	
Db	488	GGCACCAACGCTTGGACGCTGGAAGCTGCGAAGTGGCGGAGACAGATACATAAACACC	547	
Qy	481	TATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	540	
Db	548	TATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG	607	
Qy	541	GTTATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT	600	
Db	608	GTTATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT	666	
Qy	601	ACTAAGGCAAGCTTAGGTTTATATGAGCACCTTTTAGAGTGTCTTTTATGGCGCTTGTCTATGT	660	
Db	667	ACTAAGGCAAGCTTAGGTTTATATGAGCACCTTTTAGAGTGTCTTTTATGGCGCTTGTCTATGT	726	
Qy	661	TTTGTGTCGAGAGTTGTAACCATCTTGAATTAATATAAAAAAGATCATGTTTGT	717	
Db	727	TTTGTGTCGAGAGTTGTAACCATCTTGAATTAATATAAAAAAGATCATGTTTGT	783	

RESULT	5
LOCUS	AY581853
DEFINITION	Arachis hypogaea seed storage protein SSPI mRNA, partial cds.
ACCESSION	AY581853
VERSION	AY581853.1
KEYWORDS	. GI:46560479
SOURCE	Arachis hypogaea (peanut)
ORGANISM	Arachis hypogaea Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CDS

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 842)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
cDNA cloning of peanut seed storage protein
Unpublished
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
Guangdong 510275, China
FEATURES
source
1..842
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="SanYou523"
/db_xref="taxon:3818"
<1..-564
/codon_start=1
/product="seed storage protein SSP1"
/protein_id="AA00598.1"
/db_xref="GI:46560480"
/translation="DHTLFNTHSIITTTTMAKLITLVALFLAAHASARQWELQ
DRRCQSLERANLRPCEQHLMOKIQORDEDSYGRDPSPQDPSQDPRDRPSPS
PYDRRGAGSQHOCRCNELNEFNORCMCEALQIMENQSDRLQGRQEQFKREL
RNLPQQGLRAPRCLEVESGDRY"
ORIGIN
Query Match 89.9%; Score 644.6; DB 8; Length 842;
Best Local Similarity 93.9%; Pred. No. 4.5e-153;
Matches 707; Conservative 0; Mismatches 9; Indels 37; Gaps 2;
Qy 1 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCGTGCCTCCACGCGATCTCGAGGCA 60
Db 54 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCGTGCCTCCACGCGATCTCGAGGCA 113
Qy 61 GCAGTGGAACTCCAGGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 120
Db 114 GCAGTGGAACTCCAGGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 173
Qy 121 GCCTCGGAGCAACATCTCATGCAGAAATCCAAACGTCGAGGATTCATATGAACGGGA 180
Db 174 GCCTCGGAGCAACATCTCATGCAGAAATCCAAACGTCGAGGATTCATATGAACGGGA 233
Qy 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGG 240
Db 594 AAAAGAAAATAGCTTATATAGCTTATATCTATGTTATGTTTGGTATATAT 652

Qy 565 AAAGATCATCTATATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATG 624
Db 653 AAAGATCATCTATATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATG 712
Qy 625 AGCACCTTTAGAGTGCTTTTATGGCGTGTCTATGTTTGTCTGCACAGTTGTAACCA 684
Db 713 AGCACCTTTAGAGTGCTTTTATGGCGTGTCTATGTTTGTCTGCACAGTTGTAACCA 772
Qy 685 TCTTGAATAATAATAAAGATCATGTTTGT 717
Db 773 TCTTGAATAATAATAAAGATCATGTTTGT 805
RESULT 6
AY581854 678 bp mRNA linear PLN 01-MAY-2004
LOCUS Arachis hypogaea seed storage protein SSP2 mRNA, partial cds.
DEFINITION AY581854
ACCESSION AY581854
VERSION AY581854.1 GI:46560481
KEYWORDS Arachis hypogaea (peanut)
SOURCE Arachis hypogaea
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 678)
Yan, Y., Wang, L., Liao, B., Li, H., Lin, X. and Huang, S.
cDNA cloning of peanut seed storage protein
Unpublished
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (01-MAR-2004) Department of Biology and Biotechnology,
School of Life Science, Sun Yat-sen University, Guangzhou,
Guangdong 510275, China
FEATURES
Location/Qualifiers
source
1..678
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="SanYou523"
/db_xref="taxon:3818"
<1..-528
/codon_start=1
/product="seed storage protein SSP2"
/protein_id="AA00599.1"
/db_xref="GI:46560482"
/translation="DHTLFNTHSIITTTTMAKLITLVALFLAAHASARQWELQ
DRRCQSLERANLRPCEQHLMOKIQORDEDSYGRDPSPQDPSQDPSYDRRGAGSQH
QERCCNELNEFNORCMCEALQIMENQSDRLQGRQEQFKRELRLNPQQGLRAP
QRCDDLVSSEGRDY"
CDS
Qy 1 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCGTGCCTCCACGCGATCTCGAGGCA 60
Db 54 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCGTGCCTCCACGCGATCTCGAGGCA 113
Qy 61 GCAGTGGAACTCCAGGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 120
Db 114 GCAGTGGAACTCCAGGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 173
Qy 121 GCCCTCGGAGCAACATCTCATGCAGAAATCCAAACGTCGAGGATTCATATGAACGGGA 180
Db 174 GCCCTCGGAGCAACATCTCATGCAGAAATCCAAACGTCGAGGATTCATATGAACGGGA 233
Qy 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGG 240
Db 234 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGG 293

QY 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAACCAACAAAG 300
DB 294 ATCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAACCAACAAAG 353
QY 301 GTGCATGTGCGAGGACATTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCAAGGGGAG 360
DB 354 GTGCATGTGCGAGGACATTGCAACAGATCATGGAGAACCCAGAGCGATAGTTGCAAGGGGAG 413
QY 361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 420
DB 414 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 473
QY 421 GGCACACAGGCTTGCGACTTGGACCTGCAAGTGGCGGGCAGACAGATCTAAACACC 480
DB 474 GGCACACAGGCTTGCGACTTGGACCTGCAAGTGGCGGGCAGACAGATCTAAACACC 533
QY 481 TATCTCAAAAAAGAAAAAGAAAAAGAAAAAG-CTTATATATAAGCTATTATCTAT 539
DB 534 TATCTCAAAAAAGAAAAAGAAAAAGAAAAAG-CTTATATATAAGCTATTATCTAT 593
QY 540 GGTATCTTTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTGATCGGT 599
DB 594 GGTATG-TTAGTTTGGTAAATAAAGATCATCACTATATGAATGTGTGATCGGT 652
QY 600 AACTAAGCAAGCTTAGTTTATGA 625
DB 653 AACTAAGCAAGCTTAGTTTATGA 678

RESULT 7
BD107899
LOCUS BD107899 474 bp DNA linear PAT 18-SEP-2002
DEFINITION Methods and reagents for decreasing allergic reactions.
ACCESSION BD107899
VERSION BD107899.1 GI:23202717
KEYWORDS JP 2002501748-A/2.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 474)
Sosin, H., Banon, G.A., Jr, W.A.B. and Samphson, H.A.
Methods and reagents for decreasing allergic reactions
Patent: JP 200301748-A 2 22-JAN-2002;
UNIVERSITY OF ARKANSAS, MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY
UNIVERSITY OF NEW YORK, HOWARD SOSIN
OS Arachis L. (Peanut)
PN JP 2002501748-A/2
PD 22-JAN-2002
PF 29-JAN-1999 JP 2000529437
PR 31-JAN-1998 US 60/073283,13-FEB-1998 US 60/074590 PR
13-FEB-1998 US 60/074624,13-FEB-1998 US 60/074633 PR
27-AUG-1998 US 09/141220
PI HAWADO SOSIN, GARY A. BANON, WESLEY A. BIRX JR, HYU A. SAMPHON PC
C12N15/09, A01H5/00, A01K67/027, A61K39/35, C07K14/37, C07K14/415, PC
C07K14/435,
PC C07K14/47, C12N1/15, C12N1/19, C12N1/21, C12N5/00, C12N15/00, C12N5/
00
CC Methods and reagents for decreasing allergic reactions FH
Key
FT source 1..474
FT Location/Qualifiers

FEATURES
source
FT 1..474
FT Location/Qualifiers
/organism="Arachis L. (Peanut)";
/mol_type="genomic DNA";
/db_xref="taxon:32644"

ORIGIN
Query Match 66.1%; Score 474; DB 6; Length 474;
Best Local Similarity 100.0%; Pred. No. 1.1e-109;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACCGCATCTCGAGGCAG 61
DB 1 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACCGCATCTCGAGGCAG 60
QY 62 CAGTGGGAACCTCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 121
DB 61 CAGTGGGAACCTCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 120
QY 122 CCTGCGAGCAACATCTCATGCAAGAGATCCAAAGTGAAGAGGATTCATATGAACGGGAC 181
DB 121 CCTGCGAGCAACATCTCATGCAAGAGATCCAAAGTGAAGAGGATTCATATGAACGGGAC 180
QY 182 CCTGAGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGA 241
DB 181 CCTGAGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTGA 240
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAGG 301
DB 241 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAAACCAACAGG 300
QY 302 TGATGTGCGAGGATTCGAAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGGAGG 361
DB 301 TGATGTGCGAGGATTCGAAACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGGAGG 360
QY 362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
DB 361 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 420
QY 422 GCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGGCAGAGACAGATACTAA 475
DB 421 GCACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGGCAGAGACAGATACTAA 474

RESULT 8
AX155333
LOCUS AX155333 474 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 6 from Patent WO0140264.
ACCESSION AX155333
VERSION AX155333.1 GI:14536768
KEYWORDS Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea

REFERENCE 1
Bannon, G.A., Burks, W.A., Caplan, M.J., Sampson, H. and Sosin, H.
Peptide antigens
Patent: WO 0140264-A 6 07-JUN-2001;
Panacea Pharmaceuticals, LLC (US); The University of Arkansas (US)
; MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY UNIVERSITY OF NEW YORK
(US)
Location/Qualifiers
1..474
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"

FEATURES
source
FT 1..474
FT Location/Qualifiers
/organism="Arachis hypogaea";
/mol_type="unassigned DNA";
/db_xref="taxon:3818"

ORIGIN
Query Match 65.9%; Score 472.4; DB 6; Length 474;
Best Local Similarity 99.8%; Pred. No. 2.9e-109;
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACCGCATCTCGAGGCAG 61
DB 1 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACCGCATCTCGAGGCAG 60
QY 62 CAGTGGGAACCTCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 121
DB 61 CAGTGGGAACCTCAAGAGAGACAGATGCGAGAGCGAGCTCGAGAGGGGGAACCTGAGG 120
QY 122 CCTGCGAGCAACATCTCATGCAAGAGATCCAAAGTGAAGAGGATTCATATGAACGGGAC 181

Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 60
Db 11 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 70

Qy 61 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 120
Db 71 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 130

Qy 121 GCCCTGGAGCAACATCTCATGTCAGAGAAGATCCAAAGTGCAGGAGATTCATATGAACGGGA 180
Db 131 GCCCTGGAGCAACATCTCATGTCAGAGAAGATCCAAAGTGCAGGAGATTCATATGAACGGGA 190

Qy 181 CCCGTACAGCCCTAGTCAGAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 240
Db 191 CCCGTACAGCCCTAGTCAGAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 250

Qy 241 ATCTCTCAGCACCAGAGAGGTGTCGAATGAGCTGAACGAGTTTCAGAACACCAAG 300
Db 251 ATCTCTCAGCACCAGAGAGGTGTCGAATGAGCTGAACGAGTTTCAGAACACCAAG 310

Qy 301 GTGCATGTGCGAGGCAATTCGAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 360
Db 311 GTGCATGTGCGAGGCAATTCGAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 370

Qy 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTCGGGCCCTTAG 420
Db 371 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTCGGGCCCTTAG 430

Qy 421 GGCAACACAGCTTGCACATTCGACGTCGAAAGTGGCGGCGAG 462
Db 431 GGCAACACAGCTTGCACATTCGACGTCGAAAGTGGCGGCGAG 472

RESULT 11
AXI48740
LOCUS AXI48740 582 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 3 from Patent WO0136621.
ACCESSION AXI48740
VERSION AXI48740.1 GI:14347294
KEYWORDS Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE 1
Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
Down-regulation and silencing of allergen genes in transgenic
peanut seeds
Patent: WO 0136621-A 3 25-MAY-2001;
Alabama A & M University (US)
LOCATION/Qualifiers
1. .682
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"

FEATURES
source

ORIGIN
Query Match 54.4%; Score 462; DB 6; Length 582;
Best Local Similarity 100.0%; Pred. No. 1.3e-106;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 60
Db 67 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 126

Qy 61 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 120
Db 127 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 186

Qy 121 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA 180
Db 187 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA 246

Qy 181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 240
Db 247 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGG 306

Qy 241 ATCTCTCAGCACCAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 307 ATCTCTCAGCACCAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACAAAG 366

Qy 301 GTCATGTGCGAGGCAATTCGAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 360
Db 367 GTCATGTGCGAGGCAATTCGAACAGATCATGGAGACCAAGCGATAGTTGCGAGGGGAG 426

Qy 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTCGGGCCCTTAG 420
Db 427 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTCGGGCCCTTAG 486

Qy 421 GGCAACACAGCTTGCAGCTTGGACCTCGAAAGTGGCGGCGAG 462
Db 487 GGCAACACAGCTTGCAGCTTGGACCTCGAAAGTGGCGGCGAG 528

RESULT 12
AXI48738
LOCUS AXI48738 1162 bp DNA linear PAT 08-JUN-2001
DEFINITION Sequence 1 from Patent WO0136621.
ACCESSION AXI48738
VERSION AXI48738.1 GI:14347292
KEYWORDS Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE 1
Dodo,H.W., Arntzen,C.J., Konan,K.N. and Viquez,O.M.
Down-regulation and silencing of allergen genes in transgenic
peanut seeds
Patent: WO 0136621-A 1 25-MAY-2001;
Alabama A & M University (US)
LOCATION/Qualifiers
1. .1162
/organism="Arachis hypogaea"
/mol_type="unassigned DNA"
/db_xref="taxon:3818"
110. .733
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41202.1"
/db_xref="GI:14347293"
/translation="MAKLILVALALFLAAHASARQWELQDRRCQSLERANLRP
CEQHLMKIQRDEDSYERDFSPQDFSPSYDRRGAGSSQHQERCCNELNFEFENQ
RCMCEALQIMENQSDRLQROBOQFKELRNLPQCGLRAPQRCDLQVDSGRRPR
IPPILTGSRRRHQSPYGNRRYSAMCLLPRAADGDGWPFPVAVDCSG"

FEATURES
source

CDS
110. .733
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC41202.1"
/db_xref="GI:14347293"
/translation="MAKLILVALALFLAAHASARQWELQDRRCQSLERANLRP
CEQHLMKIQRDEDSYERDFSPQDFSPSYDRRGAGSSQHQERCCNELNFEFENQ
RCMCEALQIMENQSDRLQROBOQFKELRNLPQCGLRAPQRCDLQVDSGRRPR
IPPILTGSRRRHQSPYGNRRYSAMCLLPRAADGDGWPFPVAVDCSG"

ORIGIN
Query Match 54.4%; Score 462; DB 6; Length 1162;
Best Local Similarity 100.0%; Pred. No. 1.4e-106;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 60
Db 118 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCGACGAGATTCGGAGGCA 177

Qy 61 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 120
Db 178 GCAGTGGGAACCTCAAGGAGACAGAGATGCCAGCCAGCTCCAGAGGGCGAACCTGAG 237

Qy 121 GCCTCGGAGCAACATCTCATGTCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA 180

Db	238	GCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA	297
Qy	181	CCGCTACAGCCCTAGTCAGATCCGTACACCCCTAGTCCCATATGATCGGAGAGCCCTGG	240
Db	298	CCGCTACAGCCCTAGTCAGATCCGTACACCCCTAGTCCCATATGATCGGAGAGCCCTGG	357
Qy	241	ATCTCTTCAGACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAAGAACCAACAAAG	300
Db	358	ATCTCTTCAGACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAAGAACCAACAAAG	417
Qy	301	GTGATGTGCGAGCATTCGACAGATTCATGGAACACGAGCGATAGGTTGCGAGGGAG	360
Db	418	GTGATGTGCGAGCATTCGACAGATTCATGGAACACGAGCGATAGGTTGCGAGGGAG	477
Qy	361	GCACACGAGCAACAGTTCAAGAGGAGCTCAGGAATCTGCCTCAACAGTGGCGCCTTAG	420
Db	478	GCACACGAGCAACAGTTCAAGAGGAGCTCAGGAATCTGCCTCAACAGTGGCGCCTTAG	537
Qy	421	GGCACCAACAGCTTGCAGCTTGGACCTCGAAAGTGGCGGCGAG	462
Db	538	GGCACCAACAGCTTGCAGCTTGGACCTCGAAAGTGGCGGCGAG	579
RESULT 13			
LOCUS	AY158467	519 bp mRNA linear	PLN 08-DEC-2002
DEFINITION	Arachis hypogaea allergen Ara h 2.02 mRNA, complete cds.		
ACCESSION	AY158467		
VERSION	AY158467.1	GI:26245446	
KEYWORDS			
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
REFERENCE	1 (bases 1 to 519)		
AUTHORS	Chatel,J.-M., Bernard,H. and Orson,F.M.		
TITLE	cDNA of peanut allergen Ara h2 isoform		
REFERENCE	2 (bases 1 to 519)		
AUTHORS	Chatel,J.-M., Bernard,H. and Orson,F.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-OCT-2002) LIAA, INRA, CEA, CE Saclay, DRM-SPT, Gif sur Yvette 91191, France		
FEATURES	Location/Qualifiers		
source	1..519		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/db_xref="taxon:3818"		
CDS	1..519		
	/note="Ara h2 isoform"		
	/codon_start=1		
	/product="allergen Ara h 2.02"		
	/protein_id="AAM77576.1"		
	/db_xref="GI:26245447"		
	/translation="MAKLITLVALFLILAAHASARQWELQDRRCQSLERANLRP		
	CEQHLQKIQDRDSYGRDPYSPQDPDRDPYSPYSPYDRGAGSSQHOER		
	CNKLNEFENQRCMCEALQIIMENQSLRQGRQEQQFKRELNLPPQCGLRAPQRC		
	DLEVESGGDRY"		
ORIGIN			
Query Match	58.5%;	Score 419.4;	DB 8; Length 519;
Best Local Similarity	91.8%;	Pred. No. 9e-96;	
Matches 469;	Conservative 0;	Mismatches 6;	Indels 36; Gaps 1;
Qy	1	GCTCACCATACTAGTACCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGGAGGCA	60
Db	9	GCTCACCATACTAGTACCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGGAGGCA	68
Qy	61	GCATGTGGAACTCCAAGGACAGAGATGCCAGCGCTCGAGAGGGCGAACCCTGAG	120
Db	69	GCAGTGGGAATCCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG	128
Qy	121	GCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA	180
Db	129	GCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGATTCATATGAACGGGA	188
Qy	181	CCGCTACAGCCCTAGTCAG------GATCC	204
Db	189	CCGCTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCAGGACCCCGGACAGACGCTGATCC	248
Qy	205	GTACAGCCCTAGTCATATGATCGAGAGGCGCTGGATCCTCTCAACCAAGAGAGGTG	264
Db	249	GTACAGCCCTAGTCATATGATCGAGAGGCGCTGGATCCTCTCAACCAAGAGAGGTG	308
Qy	265	TTGCAATGAGCTGAACAGGTTTGAAGAACCAACAAAGGTGCATGTGCGAGGCAATTCGAACA	324
Db	309	TTGCAATGAGCTGAACAGGTTTGAAGAACCAACAAAGGTGCATGTGCGAGGCAATTCGAACA	368
Qy	325	GATCATGGAGAACACGAGCGATAGGTTGAGGGGAGGCAACAGGAGCAACAGTTCAAGAG	384
Db	369	GATAATGGAGAACACGAGCGATAGGTTGAGGGGAGGCAACAGGAGCAACAGTTCAAGAG	428
Qy	385	GGAGCTCAGGAATCTGCCTCAACAGTGGCGCCTTAGGGCACACAGCGTTGCGACTTGA	444
Db	429	GGAGCTCAGGAATCTGCCTCAACAGTGGCGCCTTAGGGCACACAGCGTTGCGACTTGA	488
Qy	445	CGTCGAAAGTGGCGGCGAGACAGATACTAA	475
Db	489	AGTCGAAAGTGGCGGCGAGACAGATACTAA	519
RESULT 14			
LOCUS	AY117434	510 bp mRNA linear	PLN 01-JUN-2003
DEFINITION	Arachis hypogaea allergen Ara h 2 isoform mRNA, partial cds.		
ACCESSION	AY117434		
VERSION	AY117434.1	GI:31322016	
KEYWORDS			
SOURCE	Arachis hypogaea (peanut)		
ORGANISM	Arachis hypogaea		
REFERENCE	1 (bases 1 to 510)		
AUTHORS	Becker,W.-M., Suhr,M., Lindner,B., Wicklein,D. and Lepp,U.		
TITLE	Re-investigation of the major peanut allergen arah2 on the molecular level		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 510)		
AUTHORS	Becker,W.-M. and Suhr,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JUN-2002) B&M Allergology, Research Center Borstel, Parkallee 35, Borstel 23845, Germany		
FEATURES	Location/Qualifiers		
source	1..510		
	/organism="Arachis hypogaea"		
	/mol_type="mRNA"		
	/db_xref="taxon:3818"		
CDS	<1..510		
	/note="conglutin"		
	/codon_start=1		
	/product="allergen Ara h 2 isoform"		
	/protein_id="AAM78596.1"		
	/db_xref="GI:31322017"		
	/translation="LTLIVAPALFLILAAHASARQWELQDRRCQSLERANLRPCEQ		
	HLMKIQDRDSYGRDPYSPQDPDRDPYSPYSPYDRGAGSSQHOERCCN		
	ELNEFENQRCMCEALQIIMENQSLRQGRQEQQFKRELNLPPQCGLRAPQRCLE		
	VESGGRDRY"		
ORIGIN			
Query Match	57.9%;	Score 415.2;	DB 8; Length 510;
Best Local Similarity	91.4%;	Pred. No. 1e-94;	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 268.476 Seconds
(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcaccatactagtagccc.....taaaagatcatgtttgtt 717

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseqn16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1980s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	100.0	717	2	AAT76615
2	717	100.0	717	2	AAT76614
3	474	66.1	474	2	AZ06383
4	474	66.1	474	4	AAS08538
5	474	66.1	474	8	ABX70606
6	474	66.1	474	10	ADG27517
7	462	64.4	682	4	AAF90337
8	462	64.4	1162	4	AAF90336
9	80	11.2	80	4	AAF90342
10	65.6	9.1	777	2	AAT86282
11	64	8.9	777	2	AAT86283
12	63.4	8.8	770	2	AAZ25346
13	62	8.6	62	4	AAF90343
14	59.2	8.3	477	10	ADH89268
15	59.2	8.3	477	12	ADG44003
16	54.4	7.6	468	10	ADH89266
17	54.4	7.6	468	12	ADG44001
18	51.4	7.2	723	2	AAT86281
19	51.4	7.2	5392	2	AAZ08720
20	49.8	6.9	62658	13	ABD33339

C	21	49.6	6.9	52216	4	AAH28355	Nucleotid
C	22	49.6	6.9	52216	6	ABL50307	Human mus
C	23	49.6	6.9	169739	6	ABQ88186	Human ost
	24	49	6.8	129	4	AAF87134	Lunasin c
	25	49	6.8	129	4	AAF83653	Soybean l
	26	49	6.8	9407	4	AAF83654	PTC9K-lu
	27	49	6.8	9408	4	AAF87135	Lunasin c
	28	48.2	6.7	7758	6	ABL33103	Lunasin c
	29	48	6.7	6509	6	ABL32226	Human imm
	30	47.4	6.6	97835	6	ABK84796	Human CDN
	31	45.6	6.4	2000	8	ADA71938	Rice gene
	32	45.6	6.4	6063	6	ABK28393	DNA trans
C	33	45.4	6.3	5823	6	ABK28382	DNA trans
	34	45.2	6.3	121724	6	ABQ88143	Human ost
	35	44.4	6.2	6050	10	ADB84181	Human lym
C	36	44.4	6.2	6385	4	AAS45353	Chemical
C	37	44.4	6.2	6385	6	ABK28186	DNA trans
	38	44.4	6.2	110000	2	AAV21209	04
	39	44	6.1	608	6	ABN62524	Human can
	40	43.8	6.1	7566	4	AAI07322	Human rep
	41	43.8	6.1	154068	11	ADP65423	Human seq
C	42	43.8	6.1	174448	11	ACN43946	Human gen
	43	43.6	6.1	60815	11	ACN43882	Human gen
	44	43.4	6.1	5919	6	ABL32389	Human imm
	45	43.4	6.1	5919	6	ABK31195	Signal tr

ALIGNMENTS

RESULT 1
AAT76615
ID AAT76615 standard; cDNA to mRNA; 717 BP.
XX
AC AAT76615;
DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX
DE Peanut allergen Ara hII cDNA clone.
XX
KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hII; ds.
XX
OS Arachis hypogaea; strain Florunner.
FH Key Location/Qualifiers
FT CDS 2..475
FT polyA_signal 562..567 /*tag= a
FT /*tag= b
XX
PN WO9724139-A1.
XX
PD 10-JUL-1997.
XX
PF 23-SEP-1996; 96WO-US015222.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 04-MAR-1996; 96US-00610424.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
XX
DR WPI; 1997-363453/33.
DR P-PSDB; AAW24153.
XX
PT Peanut allergens Ara hI and Ara hII - used for vaccination and in two-site monoclonal antibody based ELISA.
XX
PS Claim 31; Page 196; 354pp; English.

XX This cDNA clone codes for the major peanut allergen Ara hII (AAW22153),
 CC which contains multiple IgE-binding epitopes. It was amplified from a
 CC peanut seed cDNA library using a primer (see AAT58683) based on an
 CC isolated Ara hII peptide (see AAW24151). The sequence shows significant
 CC homology with the conglutinin family of seed storage proteins of other
 CC legumes. The gene is capable of producing a protein product in
 CC prokaryotic cells that is recognised by serum IgE from a large proportion
 CC of individuals with peanut hypersensitivity. Ara hII and Ara hI (see
 CC AAW24149-50) can be used to raise monoclonal antibodies which are used in
 CC a specific two-site MAb ELISA for the detection of Ara hI or Ara hII
 CC (claimed). IgE-binding Ara hII antigen epitopes (see AAW24188-93) may be
 CC used in vaccines to protect against allergic reactions to peanut
 CC allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to
 CC standardise OS field)

XX Sequence 717 BP; 217 A; 152 C; 184 G; 164 T; 0 U; 0 Other;

Query Match 100.0%; Score 717; DB 2; Length 717;
 Best Local Similarity 100.0%; Pred. No. 5.9e-178;
 Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTACCCCTCGCCCTTTCTCTGCTGCGCCACGATCTCGGAGGCA 60
 Db 1 GCTCACCATACTAGTACCCCTCGCCCTTTCTCTGCTGCGCCACGATCTCGGAGGCA 60
 Qy 61 GCAGTGGGAACCTCCAGGACACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
 Db 61 GCAGTGGGAACCTCCAGGACACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
 Qy 121 GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180
 Db 121 GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180
 Qy 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
 Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
 Qy 241 ATCTCTCAGCACCACAGAGAGGTGTCATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
 Db 241 ATCTCTCAGCACCACAGAGAGGTGTCATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
 Qy 301 GTGCATGTGCGAGGCAATTCACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAG 360
 Db 301 GTGCATGTGCGAGGCAATTCACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAG 360
 Qy 361 GCAACAGGACCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
 Db 361 GCAACAGGACCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
 Qy 421 GGCACCAACAGCTTGGACCTTGACGCTGGAAGTGGCGGAGAGACAGATACCTAAACACC 480
 Db 421 GGCACCAACAGCTTGGACCTTGACGCTGGAAGTGGCGGAGAGACAGATACCTAAACACC 480
 Qy 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
 Db 481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
 Qy 541 GTTATGTTTGTAGTTTGGTAATAATAAGATCATCATATGATGTTGATCGTGTAA 600
 Db 541 GTTATGTTTGTAGTTTGGTAATAATAAGATCATCATATGATGTTGATCGTGTAA 600
 Qy 601 ACTAAGGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTCTCTATGT 660
 Db 601 ACTAAGGCAAGCTTAGGTTATATGAGCACTTTAGAGTGCTTTTATGGCGTCTCTATGT 660
 Qy 661 TTTGTTGCTCAGAGTTGTAAACCATCTTGAATAATAATAATAATAATAATAATAATAATA 717
 Db 661 TTTGTTGCTCAGAGTTGTAAACCATCTTGAATAATAATAATAATAATAATAATAATAATA 717

RESULT 2
 AAT76614

ID AAT76614 standard; cDNA to mRNA; 717 BP.
 AC AAT76614;
 XX 17-OCT-2003 (revised)
 DT 29-DEC-1997 (first entry)
 XX Peanut allergen Ara hII cDNA clone P38.
 DE Peanut; seed storage protein; allergen; allergy; hypersensitivity;
 KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
 KW ELISA; analysis; Ara hII; ds.
 XX Arachis hypogaea; strain Florunner.
 XX Key Location/Qualifiers
 PH CDS 2..475
 FT polyA_signal /*tag= a
 FT /*tag= b
 XX WO9724139-A1.
 PN 10-JUL-1997.
 XX 23-SEP-1996; 96WO-US015222.
 PF 29-DEC-1995; 95US-0009455P.
 PR 04-MAR-1996; 96US-00610424.
 XX (UYAR-) UNIV ARKANSAS.
 PA Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
 PI WPI; 1997-363453/33.
 XX Peanut allergens Ara hI and Ara hII - used for vaccination and in two-
 PT site monoclonal antibody based ELISA.
 PS Claim 31; Page 183-185; 354pp; English.
 CC This cDNA clone, designated p38, codes for the major peanut allergen Ara
 CC hII (AAW22164), which contains multiple IgE-binding epitopes. It was
 CC amplified from a peanut seed cDNA library using a primer (see AAT58683)
 CC based on an isolated Ara hII peptide (see AAW24151). The sequence shows
 CC significant homology with the conglutinin family of seed storage proteins
 CC of other legumes. The gene is capable of producing a protein product in
 CC prokaryotic cells that is recognised by serum IgE from a large proportion
 CC of individuals with peanut hypersensitivity. Ara hII and Ara hI (see
 CC AAW24149-50) can be used to raise monoclonal antibodies which are used in
 CC a specific two-site MAb ELISA for the detection of Ara hI or Ara hII
 CC (claimed). IgE-binding Ara hII antigen epitopes (see AAW24188-93) may be
 CC used in vaccines to protect against allergic reactions to peanut
 CC allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX Sequence 717 BP; 217 A; 152 C; 184 G; 164 T; 0 U; 0 Other;

Query Match 100.0%; Score 717; DB 2; Length 717;
 Best Local Similarity 100.0%; Pred. No. 5.9e-178;
 Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCTCACCATACTAGTACCCCTCGCCCTTTCTCTGCTGCGCCACGATCTCGGAGGCA 60
 Db 1 GCTCACCATACTAGTACCCCTCGCCCTTTCTCTGCTGCGCCACGATCTCGGAGGCA 60
 Qy 61 GCAGTGGGAACCTCCAGGACACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
 Db 61 GCAGTGGGAACCTCCAGGACACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
 Qy 121 GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180
 Db 121 GCCCTGCGAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180

181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGG 240
181 CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGG 240
241 ATCTCTCAGCAACCAAGAGGCTGTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
241 ATCTCTCAGCAACCAAGAGGCTGTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
301 GTGATGTGCGAGGCTGTCGAACAGATCATGAGAACCAAGGATAGTTGCAAGGGAG 360
301 GTGATGTGCGAGGCTGTCGAACAGATCATGAGAACCAAGGATAGTTGCAAGGGAG 360
361 GCACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTGCGGCTTAG 420
361 GCACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCTCAACAGTGCGGCTTAG 420
421 GGCAACACAGCGTTGCGACCTTGGACGTCGAAAGTGGCGGAGAGACAGATCTAAACACC 480
421 GGCAACACAGCGTTGCGACCTTGGACGTCGAAAGTGGCGGAGAGACAGATCTAAACACC 480
481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
481 TATCTCAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
541 GTTATGTTTATGTTTGTGTAATAATAAAGATCATCACTATATGATGTTGATCGTTA 600
541 GTTATGTTTATGTTTGTGTAATAATAAAGATCATCACTATATGATGTTGATCGTTA 600
601 ACTAAGGCAAGCTTAGGTTATATGAGCACTTTTATGAGTGCTTTTATGCGGTTGCTATGT 660
601 ACTAAGGCAAGCTTAGGTTATATGAGCACTTTTATGAGTGCTTTTATGCGGTTGCTATGT 660
661 TTTGTTGCTGAGAGTTGTAACCACTTTGAAATAATAAAGATCATGTTTGT 717
661 TTTGTTGCTGAGAGTTGTAACCACTTTGAAATAATAAAGATCATGTTTGT 717

RESULT 3

AZ06383
ID AZ06383 standard; DNA; 474 BP.

AC AZ06383;

XX 17-OCT-2003 (revised)

DT 09-NOV-1999 (first entry)

XX Peanut allergen, Ara h 2.

XX allergy; immune response; transgenic; allergen; epitope;
KW immunoglobulin E; Ig E; binding site; peanut; ds.

XX Arachis hypogaea.

XX WO9938978-A1.

XX 05-AUG-1999.

XX 29-JAN-1999; 99WO-US002031.

XX 31-JAN-1998; 98US-0073283P.

PR 13-FEB-1998; 98US-0074590P.

PR 13-FEB-1998; 98US-0074624P.

PR 13-FEB-1998; 98US-0074633P.

PR 27-AUG-1998; 98US-00141220.

XX (UVAR-) UNIV ARKANSAS.

PA (UTNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

PA (SOSI/) SOSIN H.

XX Sosin H, Bannon GA, Burke AW, Sampson HA;

XX WPI; 1999-479189/40.

DR P-PSDB; AAY15245.

XX Modified allergen with reduced IgE binding, useful for treating e.g.
PT allergies.

XX Disclosure; Page 37; 46pp; English.

XX This is the nucleotide sequence of the Ara h 2 protein from Arachis
CC hypogaea. The Ara h 2 protein has 10 IgE (immunoglobulin E) binding
CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
CC AAY15276). By modifying the IgE binding sites the ability of the allergen
CC to provoke an immune response is downregulated. The epitopes of the IgE
CC binding sites can therefore be modified in genetically engineered plants
CC and animals to elicit less of an allergic response. (Updated on 17-OCT-
XX 2003 to standardise OS field)

SQ Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;

Query Match 66.1%; Score 474; DB 2; Length 474;

Best Local Similarity 100.0%; Pred. No. 3.3e-114; Mismatches 0; Gaps 0;

Matches 474; Conservative 0; Indels 0; Gaps 0;

QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGATCTGCGAGGAG 61

DB 1 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGATCTGCGAGGAG 60

QY 62 CAGTGGGAATCCAAAGAGACAGAAATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 121

DB 61 CAGTGGGAATCCAAAGAGAGACAGAAATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120

QY 122 CCTGCGAGCAACATCTCATGCAAGATCCACGTCGAGGAGTTTCATATGAACGGGAC 181

DB 121 CCTGCGAGCAACATCTCATGCAAGATCCACGTCGAGGAGTTTCATATGAACGGGAC 180

QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGAGCGCTGGA 241

DB 181 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATCATCGAGAGCGCTGGA 240

QY 242 TCCTCTCAGCAACCAAGAGAGGTTGCAATGAGCTGAAAGAGTTTGAAGAACCAACAAAG 301

DB 241 TCCTCTCAGCAACCAAGAGAGGTTGCAATGAGCTGAAAGAGTTTGAAGAACCAACAAAG 300

QY 302 TGCATGTGCGAGCATTGCCAAGATCATGAGAACCAAGAGCCATAGTTGCGAGGGGAGG 361

DB 301 TGCATGTGCGAGCATTGCCAAGATCATGAGAACCAAGAGCCATAGTTGCGAGGGGAGG 360

QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGCCCTTAGG 421

DB 361 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAATTCCTCAACAGTGCAGCCCTTAGG 420

QY 422 GCACCAACAGGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACATAA 475

DB 421 GCACCAACAGGTTGCGACTTTGGACGTCGAAAGTGGCGGAGAGACAGATACATAA 474

RESULT 4

AAS08538

ID AAS08538 standard; cDNA; 474 BP.

AC AAS08538;

XX 23-OCT-2001 (first entry)

XX DNA encoding anaphylactic antigen Ara h 2.

XX Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse; ss.

XX Mus sp.

XX Location/Qualifiers

XX 1..474 a

XX /*tag= a

FT	WO200140264-A2.	/product= "Peptide antigen Ara h 2"	
XX			
PN			
XX			
XX			
PD	07-JUN-2001.		
XX			
PF	06-DEC-2000; 2000WO-US033124.		
XX			
PR	06-DEC-1999; 99US-00455294.		
PR	23-JUN-2000; 2000US-0213765P.		
PR	27-SEP-2000; 2000US-0235797P.		
XX			
PA	(PANA-) PANACEA PHARM LLC.		
PA	(UYAR-) UNIV ARKANSAS.		
PA	(MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.		
XX			
PI	Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;		
XX			
DR	WPI; 2001-381378/40.		
DR	P-PSDB; AAU04707.		
XX			
PT	Antigenic fragments useful for reducing anaphylactic risk and reducing		
PT	the severity and/or number of allergic symptoms in individuals sensitive		
PT	to antigens, have reduced ability to bind immunoglobulin E.		
XX			
PS	Disclosure; Fig 10; 100pp; English.		
XX			
CC	The sequence represents the coding sequence of anaphylactic antigen Ara h		
CC	2. Ara h 2 is an anaphylactic antigen (A), which was used to design		
CC	antigenic peptides having a reduced ability to bind immunoglobulin E		
CC	(Ige) as compared with the intact (A), or having a sequence substantially		
CC	identical to a portion of sequence of an antigen that includes at least		
CC	one Ige binding site, where at least one Ige binding site of the peptide		
CC	is altered. The antigenic peptides are used in a composition which is		
CC	useful for reducing risk or severity of allergic reaction to an antigen.		
CC	This is done by identifying an individual at risk of allergic reaction to		
CC	an antigen by identifying prior display of allergic symptoms when exposed		
CC	to the antigen, or a familial relationship with an individual who		
CC	previously displayed allergic symptoms when exposed to the antigen.		
CC	Following this an antigen-specific Ige present on one or more mast cells		
CC	or basophils in the individual's serum is identified. The individual is		
CC	then contacted with a peptide corresponding to a portion of the antigen,		
CC	which is selected, formulated, and delivered so that binding of the		
CC	peptide to antigen-specific Ige is reduced as compared with Ige binding		
CC	of intact antigen. The composition is also useful for treating and		
CC	preventing allergic reactions		
XX			
SQ	Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;		
	Query Match 66.1%; Score 474; DB 4; Length 474;		
	Best Local Similarity 100.0%; Pred. No. 3.3e-114;		
	Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTCCACGCAATCTGGAGGAG 61		
Db	1 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTCCACGCAATCTGGAGGAG 60		
QY	62 CAGTGGAACTCCAAGGAGACAGAGATGCCAGGCGAGCTCCAGAGGGCGACCTGAGG 121		
Db	61 CAGTGGAACTCCAAGGAGACAGAGATGCCAGGCGAGCTCCAGAGGGCGAACTGAGG 120		
QY	122 CCCTGCGAGCAACATCTCATGCGAGAGATCCACAGCTGACGAGGATTCATATGAACGGGAC 181		
Db	121 CCCTGCGAGCAACATCTCATGCGAGAGATCCACAGCTGACGAGGATTCATATGAACGGGAC 180		
QY	182 CCGTACAGCCCTAGTCAGATTCGGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241		
Db	181 CCGTACAGCCCTAGTCAGATTCGGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 240		
QY	242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTCAGACACCAACGAG 301		
Db	241 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTCAGACACCAACGAGG 300		
QY	302 TGCATGTGGAGGCATTTGCAACAGATCATGAGAACACAGAGCATAGTTTGCAGGGGAGG 361		
Db	301 TGCATGTGGAGGCATTTGCAACAGATCATGAGAACACAGAGCATAGTTTGCAGGGGAGG 360		
QY	362 CAACAGGAGCAACAGTTTCAAGAGGGAGCTTCAGGAACCTTGCCTCAACAGTGCAGCCCTTAGG 421		
Db	361 CAACAGGAGCAACAGTTTCAAGAGGGAGCTTCAGGAACCTTGCCTCAACAGTGCAGCCCTTAGG 420		
QY	422 GCACACACAGCTTGGCAGCTTGGACGTCGAAAGTGGCGGCAGAGACAGATACTAA 475		
Db	421 GCACACACAGCTTGGCAGCTTGGACGTCGAAAGTGGCGGCAGAGACAGATACTAA 474		
	RESULT 5		
	ABX70606		
ID	ABX70606 standard; cDNA; 474 BP.		
XX			
AC	ABX70606;		
XX			
DT	26-MAR-2003 (first entry)		
XX			
DE	Peanut Ara h2 cDNA.		
XX			
KW	Peanut; ss; allergy; Ara h1; Ara h2; Ara h3; gene; IgE binding site;		
KW	anaphylactic food allergen; anti-allergenic; vaccine; wound healing.		
XX			
OS	Arachis hypogaea.		
XX			
PN	WO200274250-A2.		
XX			
PD	26-SEP-2002.		
XX			
PF	18-MAR-2002; 2002WO-US009108.		
XX			
PR	16-MAR-2001; 2001US-0276822P.		
PR	18-MAR-2002; 2002US-00276822.		
XX			
PA	(PANA-) PANACEA PHARM.		
XX			
PI	Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;		
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
PI	Rabjohn PA, Shin DS, Stanley JS;		
XX			
DR	WPI; 2003-018765/01.		
DR	P-PSDB; ABUS2464.		
XX			
PT	New modified anaphylactic food allergen, useful for preventing or		
PT	treating allergic reactions associated with e.g. anaphylactic allergens.		
XX			
PS	Example 10; Fig 40; 300pp; English.		
XX			
CC	The invention relates to a modified anaphylactic food allergen has an		
CC	amino acid sequence that is substantially identical to that of natural		
CC	anaphylactic food allergen, except for a cysteine residue that has been		
CC	modified so that it cannot participate in the disulphide bond. The		
CC	modification may also comprise mutation of the Ige binding sites to		
CC	reduce allergenicity. Also included are: (1) a method of making a		
CC	modified anaphylactic food allergen; (2) a nucleotide molecule encoding		
CC	or for causing a site specific mutation in the modified anaphylactic food		
CC	allergen; (3) a transgenic plant or animal expressing the modified		
CC	anaphylactic food allergen; (4) a method of treating an individual by		
CC	reducing the clinical response to a natural anaphylactic food allergen;		
CC	and an isolated fragment of peanut allergen Ara h 1. The modified		
CC	anaphylactic food allergen is useful for preventing or treating allergic		
CC	reactions associated with any natural allergen such as food, insect,		
CC	rubber or preferably anaphylactic allergens. It is also useful for		
CC	treating wounds in mammals such as bovine, canine, feline, caprine,		
CC	ovine, porcine, murine or equine species. The present sequence is a cDNA		
CC	encoding a peanut allergen (e.g. Ara h1, h2 or h3)		
XX			
SQ	Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;		
	Query Match 66.1%; Score 474; DB 8; Length 474;		

Best Local Similarity 100.0%; Pred. No. 3.3e-114;		Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGAGGACG 61	
Db	1	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGAGGACG 60	
Qy	62	CAGTGGGAATCCCAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAGG 121	
Db	61	CAGTGGGAATCCCAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAGG 120	
Qy	122	CCCTGCGAGCAACATCTCATGCAAGATCCAAAGTGCAGAGGATTCATATGAACGGGAC 181	
Db	121	CCCTGCGAGCAACATCTCATGCAAGATCCAAAGTGCAGAGGATTCATATGAACGGGAC 180	
Qy	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGGA 241	
Db	181	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGGCGCTGGA 240	
Qy	242	TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTTGAGAACCAACCAAGG 301	
Db	241	TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTTGAGAACCAACCAAGG 300	
Qy	302	TGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGCGATAGTTGCAGGGGAGG 361	
Db	301	TGCATGTGCGAGGCAATTCGAACAGATCATGAGAACCAAGAGCGATAGTTGCAGGGGAGG 360	
Qy	362	CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCTCAACAGTGGCGCTTTAGG 421	
Db	361	CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCTCAACAGTGGCGCTTTAGG 420	
Qy	422	GCACCAAGGCTTGGAGCTGGAGCTGGAAGTGGCGGCGAGAGACAGATCTAA 475	
Db	421	GCACCAAGGCTTGGAGCTTGGAGCTGGAAGTGGCGGCGAGAGACAGATCTAA 474	
RESULT 6			
ADG27517			
ID	ADG27517 standard; cDNA; 474 BP.		
XX	AC ADG27517;		
XX	26-FEB-2004 (first entry)		
XX	Peanut cDNA encoding Ara h2.		
XX	Peanut; plant; ss; gene; allergen; Ara h1; Ara h2; Ara h3;		
KW	Glycinin A2B1a; Jug nl; antiallergic; vulnerary;		
KW	anaphylactic food allergen; IgE; allergy; wound.		
XX	Arachis hypogaea.		
XX	US2003202980-A1.		
XX	30-OCT-2003.		
XX	18-MAR-2002; 2002US-00100303.		
XX	29-DEC-1995; 95US-0009455P.		
PR	23-SEP-1996; 96US-000717933.		
PR	31-JAN-1998; 98US-0073283P.		
PR	13-FEB-1998; 98US-0074590P.		
PR	13-FEB-1998; 98US-0074633P.		
PR	29-JUN-1998; 98US-00106872.		
PR	27-AUG-1998; 98US-00141220.		
PR	13-NOV-1998; 98US-00191593.		
PR	29-JAN-1999; 99US-00240557.		
PR	29-JAN-1999; 99US-00241101.		
PR	11-FEB-1999; 99US-00248673.		
PR	11-FEB-1999; 99US-00248674.		
PR	02-MAR-1999; 99US-0122450P.		
PR	02-MAR-1999; 99US-0122452P.		
PR	02-MAR-1999; 99US-0122560P.		
PR	02-MAR-1999; 99US-0122565P.		
PR	02-MAR-1999; 99US-0122566P.		
PR	11-MAR-1999; 99US-00267719.		
PR	28-JAN-2000; 2000US-00494096.		
PR	16-MAR-2001; 2001US-0276822P.		
XX	(CAPL/) CAPLAN M J.		
PA	(SOSI/) SOSIN H B.		
PA	(SAMP/) SAMPSON H.		
PA	(BANN/) BANNON G A.		
PA	(BURK/) BURKS A W.		
PA	(COCK/) COCKRELL G.		
PA	(COMP/) COMPADRE C M.		
PA	(CONN/) CONNAUGHTON C.		
PA	(HELM/) HELM R M.		
PA	(KING/) KING N E.		
PA	(KOPP/) KOPPER R A.		
PA	(MALE/) MALEKI S J.		
PA	(RABJ/) RABJOHN P A.		
PA	(SHIN/) SHIN D S.		
PA	(STAN/) STANLEY J S.		
XX	Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;		
PI	Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;		
PI	Rabjohn PA, Shin DS, Stanley JS;		
XX	WPI; 2003-875632/81.		
DR	P-PSDB; ADG27518.		
XX	New modified anaphylactic food allergen comprising a cysteine residue		
PT	which has been modified so that it cannot participate in the disulfide		
PT	bond, useful for treating allergic reactions or wounds.		
XX	Example 10; SEQ ID NO 62; 194pp; English.		
PS	The invention relates to a modified anaphylactic food allergen whose		
CC	amino acid sequence is substantially identical to that of a natural		
CC	anaphylactic food allergen. The natural anaphylactic food allergen		
CC	includes at least one cysteine residue that participates in a disulphide		
CC	bond when the natural anaphylactic food allergen is in its native		
CC	conformation, except that the cysteine residue has been modified so that		
CC	it cannot participate in the disulphide bond. Also included are a method		
CC	of making a modified anaphylactic food allergen, a nucleotide molecule		
CC	encoding a modified anaphylactic food allergen defined above, a		
CC	nucleotide molecule for causing a site specific mutation in a gene		
CC	encoding a natural anaphylactic food allergen, a transgenic plant or		
CC	animal expressing a modified anaphylactic food allergen defined above, a		
CC	method of treating an individual by reducing the clinical response to a		
CC	natural anaphylactic food allergen by administering a modified		
CC	anaphylactic food allergen and an isolated fragment of peanut allergen		
CC	Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or		
CC	ADG27465. About 10-17% of the amino acids have been modified in at least		
CC	one IgE epitope or all the IgE epitopes recognised when the natural		
CC	anaphylactic food allergen is contacted with serum IgE from individual(s)		
CC	allergic to the natural anaphylactic food allergen. The invention		
CC	discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding		
CC	cDNAs), Soybean Glycinin A2B1a and IgE-binding epitopes of the English		
CC	walnut allergen Jug nl. The modified anaphylactic food allergen can be		
CC	used for treating allergic reactions or wounds. The present sequence		
CC	encodes a Peanut allergen of the invention.		
XX	Sequence 474 BP; 132 A; 125 C; 139 G; 78 T; 0 U; 0 Other;		
Qy	Query Match 66.1%; Score 474; DB 10; Length 474;		
Db	Best Local Similarity 100.0%; Pred. No. 3.3e-114;		
Qy	Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGAGGACG 61	
Db	1	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCCGCAGCATCTCGAGGACG 60	
Qy	62	CAGTGGGAATCCCAAGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAGG 121	

Db 61 CAGTGGGAATCTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTTGAGG 120
Qy 122 CCCTGCGAGCAACATCTCATGCAAGATCCAGGTCAACGAGGATTCATATGACGGGAC 181
Db 121 CCCTGCGAGCAACATCTCATGCAAGATCCAGGTCAACGAGGATTCATATGACGGGAC 180
Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTTGG 241
Db 181 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGGCGCTTGG 240
Qy 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 241 TCCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAACGAGTTTGAGAACACCAAGG 300
Qy 302 TGCATGTGCGGGCATTTGCACAGATCATGAGAACCCAGAGCCGATAGTTTCAGGGGAGG 361
Db 301 TGCATGTGCGGGCATTTGCACACAGATCATGAGAACCCAGAGCCGATAGTTTCAGGGGAGG 360
Qy 362 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAATCTTGCTCAACAGTGGCGCTTAGG 421
Db 361 CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAATCTTGCTCAACAGTGGCGCTTAGG 420
Qy 422 GCACCAAGCGTTTCCGACTTGGACGTTGGAAGTGGAGGCGGCGAGACAGATACTAA 475
Db 421 GCACCAAGCGTTTCCGACTTGGACGTTGGAAGTGGAGGCGGCGAGACAGATACTAA 474

RESULT 7
AAF90337
ID AAF90337 standard; DNA; 682 BP.
XX AC AAF90337;
XX AC AAF90337;
DT 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX XX
DE Peanut allergen Ara h2 gene.
XX XX
KW Peanut; allergen; Ara h2; Ara h6; Ara h7; transgenic plant; allergy; ds.
XX OS Arachis hypogaea.
XX FH Key Location/Qualifiers
XX CDS 59..682
FT /*tag= a
FT sig_peptide 59..121
FT /*tag= b
FT mat_peptide 122..679
FT /*tag= c
FT misc_feature 321..528
FT /*tag= d
FT /note= "PCR amplified fragment"
XX XX
PN WO200136621-A2.
XX PN
XX PD 25-MAY-2001.
XX PD
PF 20-NOV-2000; 2000WO-US031657.
XX PF
PR 19-NOV-1999; 99US-0167255P.
XX PR
XX (UVAL-) UNIV ALABAMA A & M.
XX PA
XX PI Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX XX
DR WPI; 2001-355630/37.
DR P-PSDB; AAB82383.
XX DR
XX PT Producing transgenic peanut plants that produce allergen-free seeds,
XX PT useful in non-allergenic foods, by antisense or sense co-suppression of
XX FT allergen-encoding genes.
XX FT

PS Claim 6; Fig 3; 72pp; English.
XX XX
CC The present sequence is that of the peanut allergen Ara h2 gene coding
CC region. A portion of the Ara h2 gene is homologous to corresponding
CC regions of the peanut allergen Ara h6 and Ara h7 genes. This region has
CC been PCR amplified, cloned in transformation vectors (pUC18 and pBI4434)
CC in sense and antisense orientations and used to down-regulate Ara h2, Ara
CC h6 and Ara h7 allergens in peanut. This is an example of the method of
CC the invention, which relates to the production of a peanut plant having
CC reduced, or undetectable, allergenic protein (AP) content in its seed. A
CC peanut plant cell is transformed with a DNA construct containing an
CC antisense AP gene and/or sense AP gene, or their fragments, regenerated
CC to plants, and fertile transgenic plants that produce seeds with reduced
CC AP content are identified. The AP sense or antisense gene may comprise at
CC least a portion of the Ara h2 gene. The seeds are useful for preparation
CC of allergen-free foods. (Updated on 11-SEP-2003 to standardise OS field)
XX XX
SQ Sequence 682 BP; 170 A; 189 C; 202 G; 121 T; 0 U; 0 Other;

Query Match 64.4%; Score 462; DB 4; Length 682;
Best Local Similarity 100.0%; Pred. No. 5,3e-111;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGTCAACATACCTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCAATCTGCGAGGCA 60
Db 67 GCTCACATACCTAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCAATCTGCGAGGCA 126
Qy 61 GCAGTGGGAATCCAGAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 120
Db 127 GCAGTGGGAATCCAGAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCCTGAG 186
Qy 121 GCCCTGCGAGCAACATCTCATGCAAGATCCAACTGACGAGGATTCATATGAACGGGA 180
Db 187 GCCCTGCGAGCAACATCTCATGCAAGATCCAACTGACGAGGATTCATATGAACGGGA 246
Qy 181 CCGGTACAGCCCTAGTCAGGATCCGTAACGCCCTAGTCCATATGATCGAGAGGGCGCTGG 240
Db 247 CCGGTACAGCCCTAGTCAGGATCCGTAACGCCCTAGTCCATATGATCGAGAGGGCGCTGG 306
Qy 241 ATCTCTCTCAGCACCAAGAGAGGTGTTCATAGCTGAACTGAGAGTTTGAGAACCAACAAAG 300
Db 307 ATCTCTCTCAGCACCAAGAGAGGTGTTCATAGCTGAACTGAGAGTTTGAGAACCAACAAAG 366
Qy 301 GTGCATGTGCGAGGCATTCACACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 360
Db 367 GTGCATGTGCGAGGCATTCACACAGATCATGGAGAACCCAGAGCGATAGTTGCGAGGGAG 426
Qy 361 GCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACTTGCTCTCAACAGTCGGCGCTTAG 420
Db 427 GCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACTTGCTCTCAACAGTCGGCGCTTAG 486
Qy 421 GGCACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 462
Db 487 GGCACACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 528

RESULT 8
AAF90336
ID AAF90336 standard; DNA; 1162 BP.
XX AC AAF90336;
XX AC AAF90336;
DT 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX XX
DE Peanut allergen Ara h2 gene.
XX XX
KW Peanut; allergen; Ara h2; transgenic plant; allergy; ds.
XX OS Arachis hypogaea.
XX FH Key Location/Qualifiers
XX TATA_signal 39..46

```

FT CDS /*tag= a
FT 110..733
FT /*tag= b
FT sig_peptide 110..172
FT /*tag= c
FT mat_peptide 172..730
FT /*tag= d
FT polyA_signal 1060..1065
FT /*tag= e
XX
XX WO200136621-A2.
XX
XX 25-MAY-2001.
XX
XX 20-NOV-2000; 2000WO-US031657.
XX
XX 19-NOV-1999; 99US-0167255P.
XX
XX (UYAL-) UNIV ALABAMA A & M.
XX
XX Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX
XX WPI; 2001-355630/37.
XX
XX P-PSDB; AAB82383.
XX
XX Producing transgenic peanut plants that produce allergen-free seeds,
XX useful in non-allergenic foods, by antisense or sense co-suppression of
XX allergen-encoding genes.
XX
XX Claim 5; Fig 2; 72pp; English.
XX
XX The present sequence is that of the peanut allergen Ara h2 gene. The gene
XX was identified following screening of genomic DNA using a probe (see
XX AAF90342) based on the known Ara h2 cDNA sequence, subcloning of positive
XX clones into a phagemid vector, restriction enzyme digestion, ligation
XX with vector DNA, Southern hybridisation and sequencing. The gene encodes
XX a 207-amino acid allergenic protein (see AAB82383). The invention relates
XX to a method for producing a peanut plant having reduced, or undetectable,
XX allergenic protein (AP) content in its seed. A peanut plant cell is
XX transformed with a DNA construct containing an antisense AP gene and/or
XX transgenic plants that produce seeds with reduced AP content are
XX identified. The AP sense or antisense gene may comprise at least a
XX portion of the present sequence. The seeds are useful for preparation of
XX allergen-free foods. Recombinant AP may be produced and used to produce
XX antibodies useful for detecting AP in foods, and for treatment or
XX prevention of peanut allergy. (Updated on 11-SEP-2003 to standardise OS
XX field)
XX
XX Sequence 1162 BP; 272 A; 325 C; 329 G; 236 T; 0 U; 0 Other;
XX
XX Query Match 64.4%; Score 462; DB 4; Length 1162;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-111;
XX Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1 GCTCACATAGTACCCCTCGCCCTTTCTCTGCTGCCACGCACTCGGAGGCA 60
Db 118 GCTCACATAGTACCCCTCGCCCTTTCTCTGCTGCCACGCACTCGGAGGCA 177
Qy 61 GCAGTGGGAAGTCCAGGAGACAGATGCCAGGCCAGCTCGAGAGGCGCAACCTGAG 120
Db 178 GCAGTGGGAAGTCCAGGAGACAGATGCCAGGCCAGCTCGAGAGGCGCAACCTGAG 237
Qy 121 GCCCTGCGAGCAACATCTCATGCGAGAAGATCCACGCTGACGAGGATTTCATATGAACGGGA 180
Db 238 GCCCTGCGAGCAACATCTCATGCGAGAAGATCCACGCTGACGAGGATTTCATATGAACGGGA 297
Qy 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 298 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 357
Qy 241 ATCTCTCAGCACCAAGAGGTTGTCATAGCTGAGCTGAAAGAGTTTGAGAACCAACCAAG 300

```

```

Db 358 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGGAGAACCAACAAAG 417
Qy 301 GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCGAGAGCATAGGTTGCGAGGAG 360
Db 418 GTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCGAGAGCATAGGTTGCGAGGAG 477
Qy 361 GCACACAGGAGCACAGCTTCAAGAGGAGGAGCTCAGGAACTTGCCTCAACAGTGCGGCCTTAG 420
Db 478 GCAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACTTGCCTCAACAGTGCGGCCTTAG 537
Qy 421 GGCAACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 462
Db 538 GGCAACACAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAG 579
XX
XX RESULT 9
XX AAF90342
XX ID AAF90342 standard; DNA; 80 BP.
XX
XX AC AAF90342;
XX
XX 11-SEP-2003 (revised)
XX 23-JUL-2001 (first entry)
XX
XX Peanut allergen Ara h2 gene probe.
XX
XX Peanut; allergen; Ara h2; transgenic plant; allergy; probe; ss.
XX
XX Arachis hypogaea.
XX
XX WO200136621-A2.
XX
XX 25-MAY-2001.
XX
XX 20-NOV-2000; 2000WO-US031657.
XX
XX 19-NOV-1999; 99US-0167255P.
XX
XX (UYAL-) UNIV ALABAMA A & M.
XX
XX Dodo HW, Arntzen CJ, Konan KN, Viquez OM;
XX
XX WPI; 2001-355630/37.
XX
XX Producing transgenic peanut plants that produce allergen-free seeds,
XX useful in non-allergenic foods, by antisense or sense co-suppression of
XX allergen-encoding genes.
XX
XX Example 1; Page 44; 72pp; English.
XX
XX The present sequence is that of an oligonucleotide probe corresponding to
XX nucleotides 11-91 of a published peanut allergen Ara h2 cDNA sequence.
XX The probe was used to screen a peanut genomic library for the isolation
XX of the Ara h2 gene (see AAF90336). The invention relates to a method for
XX producing a peanut plant having reduced, or undetectable, allergenic
XX protein (AP) content in its seed. A peanut plant cell is transformed with
XX a DNA construct containing an antisense AP gene and/or sense AP gene, or
XX their fragments, regenerated to plants, and fertile transgenic plants
XX that produce seeds with reduced AP content are identified. The AP sense
XX or antisense gene may comprise at least a portion of the present
XX sequence. The seeds are useful for preparation of allergen-free foods.
XX Recombinant AP may be produced and used to produce antibodies useful for
XX detecting AP in foods, and for treatment or prevention of peanut allergy.
XX (Updated on 11-SEP-2003 to standardise OS field)
XX
XX Sequence 80 BP; 17 A; 27 C; 21 G; 15 T; 0 U; 0 Other;
XX
XX Query Match 11.2%; Score 80; DB 4; Length 80;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-11;
XX Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 11 CTAGTAGCCCTCGCCCTTTCTCTGCTGCCACGCACTCTCGAGGAGGAGTGGGA 70

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 1681.67 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gccacacatactagtagccc.....taaaagatcatgttttgtt 717

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0*
Maximum Match 100*
Listing first 45 summaries

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_est8.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	686.4	95.7	788	6	CD038434
2	666	92.9	689	6	CD038213
3	555.2	77.4	732	6	CD038384
4	524	73.1	524	7	CD0897505
5	487.8	68.0	634	6	CD038323
6	315	43.9	405	6	CD038306
7	289	40.3	308	6	CD038459
8	283.4	39.5	375	7	CD0897504
9	223.8	31.2	696	6	CD038415
10	215.2	30.0	732	6	CD038394
11	208.8	29.1	661	6	CD038326
12	143.4	20.0	211	7	CD0897508
13	137.6	19.2	286	7	CD0897506
14	133.2	18.6	680	6	CD038092
15	131.8	18.4	351	7	CD0897507
16	129.6	18.1	714	6	CD038841
17	81	11.3	290	7	CD0897509
18	71.6	10.0	659	7	CN825259
19	67.6	9.4	305	6	CD038103
20	65.8	9.2	445	2	BF594963
21	63	8.8	719	2	BE661358
22	62.6	8.7	707	2	BE658140
23	62.4	8.7	479	4	BG237204
24	60.6	8.5	504	1	AJ498204

c

ALIGNMENTS

RESULT 1
LOCUS CD038434
DEFINITION UTPPI007 C04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION CD038434
VERSION CD038434.1 GI:30420272
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 788)
Luo M., Dang P., Guo B.Z., Holbrook C.C., Lee R.D., Bausher M.G.
and Lynch R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.

FEATURES
source

Location/Qualifiers
1..788
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI007 C04"
/tissue_type="immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent and ultrapurified (GIBCOBRL). mRNA was extracted and purified from

25 59.8 8.3 468 1 AI960869
26 59.2 8.3 469 2 AW279501
27 59.2 8.3 469 2 AW318110
28 59.2 8.3 471 2 AW318155
29 59.2 8.3 474 2 AW318193
30 59.2 8.3 475 2 AW395515
31 59.2 8.3 475 2 AW395679
32 59.2 8.3 475 2 AW398039
33 59.2 8.3 477 1 AI759696
34 59.2 8.3 479 2 BF595027
35 59.2 8.3 479 2 AW395742
36 59.2 8.3 481 2 AW396869
37 59.2 8.3 482 2 AW397141
38 59.2 8.3 486 1 AI941225
39 59.2 8.3 487 2 AW397497
40 59.2 8.3 489 2 BF596512
41 59.2 8.3 490 1 AI748087
42 59.2 8.3 490 2 AW397370
43 59.2 8.3 490 2 AW398070
44 59.2 8.3 495 2 AW397013
45 59.2 8.3 521 2 AW395825

CD038434 788 bp mRNA linear EST 07-MAY-2003
UTPPI007 C04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI007_C04 5', mRNA sequence.

CD038434

CD038434.1 GI:30420272

EST.

Arachis hypogaea (peanut)

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;

Aeschynomeneae; Arachis.

1 (bases 1 to 788)

Luo M., Dang P., Guo B.Z., Holbrook C.C., Lee R.D., Bausher M.G.

and Lynch R.E.

Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

Contact: Baozhu Guo

Molecular Genetics

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

Location/Qualifiers

1..788

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPPI007 C04"

/tissue_type="immature pods"

/dev_stage="R6"

/lab_host="XLI-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent and ultrapurified (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 95.7%; Score 686.4; DB 6; Length 788;
Best Local Similarity 99.7%; Pred. No. 4.6e-160;
Matches 698; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GCTCACATACTAGTAGCCCTCCCTTTTCCTCGCTGCGCCACGCACTGCGAGGCA 60
Db |||||
90 GCTCACATACTAGTAGCCCTCCCTTTTCCTCGCTGCGCCACGCACTGCGAGGCA 149
Qy 61 GCAGTGGGAACCTCCAGGACACAGATGCGCAGAGCCAGCTCGAGGGCGAACCTGAG 120
Db |||||
150 GCAGTGGGAACCTCCAGGACACAGATGCGCAGAGCCAGCTCGAGGGCGAACCTGAG 209
Qy 121 GCCTTCGAGCAACATCTCATGCGAGAAGATGCCAAGCTGACGAGATTCTATGAACGGGA 180
Db |||||
210 GCCTTCGAGCAACATCTCATGCGAGAAGATGCCAAGCTGACGAGATTCTATGAACGGGA 269
Qy 181 CCGGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGG 240
Db |||||
270 CCGGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGCTGG 329
Qy 241 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACACCAAG 300
Db |||||
330 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACACCAAG 389
Qy 301 GTGCATGTGCGAGGCATTGCAACAGATCATGAGAACACGAGCGATAGTTGCGAGGGGAG 360
Db |||||
390 GTGCATGTGCGAGGCATTGCAACAGATCATGAGAACACGAGCGATAGTTGCGAGGGGAG 449
Qy 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 420
Db |||||
450 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAG 509
Qy 421 GGCACACAGCGTTGCGACTTGCACTCGAAAGTGGCGGCGAGACAGATATAACACC 480
Db |||||
510 GGCACACAGCGTTGCGACTTGCACTCGAAAGTGGCGGCGAGACAGATATAACACC 569
Qy 481 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAGCTATTATCTATG 540
Db |||||
570 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAGCTATTATCTATG 629
Qy 541 GTTATGTTAGTTTGGTATAATAAAGATCATCACTATATAGTATGATCGTGTTA 600
Db |||||
630 GTTATG-TTATGTTTGGTATAATAAAGATCATCACTATATAGTATGATCGTGTTA 688
Qy 601 ACTAAGCGAAGCTTAGTTATATAGCACCTTTAGAGTCTTTTATGGCGTTGCTCTATCT 660
Db |||||
689 ACTAAGCGAAGCTTAGTTATATAGCACCTTTAGAGTCTTTTATGGCGTTGCTCTATCT 748
Qy 661 TTTGTTGCTGCGAGAGTTGTAACCATCTTGAAATAATATAA 700
Db |||||
749 TTTGTTGCTGCGAAGTTGTAACCATCTTGAAATAATATAA 788

RESULT 2

CD038213
LOCUS CD038213 689 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPII004 A09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPII004_A09 5', mRNA sequence.
ACCESSION CD038213
VERSION CD038213.1 GI:30420051

KEYWORDS

SOURCE

ORGANISM

EST.

Arachis hypogaea (peanut)

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

1 (bases 1 to 689)

Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.

and Lynch, R. E.

Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

Contact: Baozhu Guo

Molecular Genetics

USDA/ARS, Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

FEATURES

source

Location/Qualifiers

1..689

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPII004_A09"

/tissue_type="Immature pods"

/dev_stage="R6"

/lab_host="XLI-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;

cDNA library was constructed from peanut cultivar A13;

(NCV11XAR4). A13 has resistance to Aspergillus infection

and drought tolerance. The immature pods that developed to

R6 stage were collected from different plants, and placed

into liquid N2 immediately and stored in -80oC freezer.

Total RNA was isolated with TRIZOL-Reagent

ultrapure (GIBCOBRL). mRNA was extracted and purified from

total RNA (Promega). cDNA synthesis and library

construction followed the protocol of by ZAP-cDNA Gigapack

III Gold cloning kit (Stratagene). The cDNA above 500bp

were collected after size-fraction. The inserts were

directionally cloned into Uni-ZAP XR vector using XhoI

EcoRI sites adapters. The lambda library was packed into

phages using Gigapack III Gold (Stratagene). The

un-amplified library was used to excise pBluescript

phagemids from the Uni-ZAP XR vector, and the phagemids

was used to transform the host bacteria SOLR. The library

was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 92.9%; Score 666; DB 6; Length 689;
Best Local Similarity 99.6%; Pred. No. 5.4e-155;
Matches 677; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 38 GCTGCCCGCCGATCTGCGAGGCGAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGC 97
Db |||||
1 GCTGCCCGCCGATCTGCGAGGCGAGTGGGAACCTCCAGGAGACAGAGATGCCAGAGC 60
Qy 98 CAGCTCGAGAGGGCGGAACCTGAGGCCCTCGAGCAACATCTCATGACAGAGATCCAAAGT 157
Db |||||
61 CAGCTCGAGAGGGCGGAACCTGAGGCCCTCGAGCAACATCTCATGACAGAGATCCAAAGT 120
Qy 158 GACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGT 217
Db |||||
121 GACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGT 180
Qy 218 CCATATGATCGGAGAGGCGCTGGATCTCTCAGCACCAAGAGAGGCTGTTGCAATCAGCTG 277
Db |||||
181 CCATATGATCGGAGAGGCGCTGGATCTCTCAGCACCAAGAGAGGCTGTTGCAATCAGCTG 240

QY 278 AACGAGTTTGAGAACACCAAGGTGCATGTCGGAGGSCATTGCAACAGATCATGGAGAAC 337
DB 241 AACGAGTTTGAGAACACCAAGGTGCATGTCGGAGGSCATTGCAACAGATCATGGAGAAC 300
QY 338 CAGAGCGATAGTTGTCAGGGGAGGCAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAAC 397
DB 301 CAGAGCGATAGTTGTCAGGGGAGGCAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAAC 360
QY 398 TTGCCTCAACAGTCGGCCCTTAGGGACACACAGGTTGCGACTTGGACGTGCAAGTGGC 457
DB 361 TTGCCTCAACAGTCGGCCCTTAGGGACACACAGGTTTGGACTTGGACGTGCAAGTGGC 420
QY 458 GGCAGACAGACATCTAAACACCTTCTCAAAAAAGAAAGAAAGAAAGAAAGAAATAGC 517
DB 421 GGCAGACAGACATCTAAACACCTTCTCAAAAAAGAAAGAAAGAAAGAAAGAAATAGC 480
QY 518 TTATATTAAGCTATTATCTATGTTTATGTTTATGTTTGGTAAATAATAAAGATCATCACT 577
DB 481 TTATATTAAGCTATTATCTATGTTTATGTTTATGTTTGGTAAATAATAAAGATCATCACT 539
QY 578 ATATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAG 637
DB 540 ATATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAG 599
QY 638 TGCTTTTATGCGTGTCTATGTTTCTGCTGACAGAGTTGTAAACCATCTTGAATAATA 697
DB 600 TGCTTTTATGCGTGTCTATGTTTCTGCTGACAGAGTTGTAAACCATCTTGAATAATA 659
QY 698 TAAAGAGATCATGTTTGT 717
DB 660 TAAAGAGATCATGTTTGT 679

RESULT 3

CD038384

LOCUS

DEFINITION UTPI006_E02 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP) EST 07-MAY-2003

ACCESSION CD038384

VERSION CD038384.1

KEYWORDS GI:30420222

SOURCE Arachis hypogaea (peanut)

ORGANISM Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
1 (bases 1 to 732)

REFERENCE

AUTHORS Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.

and Lynch, R. E.

Generation and Analyses of ESTs for Arachis hypogaea

Unpublished (2003)

Contact: Baozhu Guo

USDA/ARS Crop Protection and Management Research Unit

2747 Davis Rd., Tifton, GA 31794, USA

Tel: 229-387-2334

Fax: 229-387-2321

Email: bguo@tifton.usda.gov

Seq primer: T3.

FEATURES

source

1. .732

/organism="Arachis hypogaea"

/mol_type="mRNA"

/cultivar="A13"

/db_xref="taxon:3818"

/clone="UTPI006_E02"

/tissue_type="Immature pods"

/dev_stage="R6"

/lab_host="X11-blue"

/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"

/note="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;

cdna library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 77.4%; Score 555.2; DB 6; Length 732;
Best Local Similarity 93.1%; Pred. No. 2.2e-127;
Matches 617; Conservative 0; Mismatches 9; Indels 37; Gaps 2;

QY 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCCGCCATCTCGGAGGCA 60
DB 70 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCCGCCATCTCGGAGGCA 129
QY 61 GCAGTGGGAATCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 120
DB 130 GCAGTGGGAATCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAG 189
QY 121 GCGCTGCGAGCAACATCTCATGCGAAGATCCAAAGTGAAGAGATTCATATGAACGGGA 180
DB 190 GCGCTGCGAGCAACATCTCATGCGAAGATCCAAAGTGAAGAGATTCATATGAACGGGA 249
QY 181 CCGGTACAGCCCTAGTCAG-----CATCC 204
DB 250 CCGGTACAGCCCTAGTCAG-----CATCC 309
QY 205 GTACAGCCCTAGTCATATGATCGGAGAGCGCTGGATCTCTCAGCAACCAAGAGAGGTG 264
DB 310 GTACAGCCCTAGTCATATGATCGGAGAGCGCTGGATCTCTCAGCAACCAAGAGAGGTG 369
QY 265 TTGCAATGAGCTGAACGAGTTTGAGAAACAAACAAAGGTGCATGTGCGAGGCATTGCAACA 324
DB 370 TTGCAATGAGCTGAACGAGTTTGAGAAACAAACAAAGGTGCATGTGCGAGGCATTGCAACA 429
QY 325 GATCATGGAGAACCAAGAGCGATAGTTGCGAGGGGCAACAGGAGCAACAGTTCAGAG 384
DB 430 GATAATGGAGAACCAAGAGCGATAGTTGCGAGGGGCAACAGGAGCAACAGTTCAGAG 489
QY 385 GGAGCTCAGGAACTTGCCCTCAACAGTGGCGCTTAGGGCAACCAAGAGGTTCGACTTGA 444
DB 490 GGAGCTCAGGAACTTGCCCTCAACAGTGGCGCTTAGGGCAACCAAGAGGTTCGACTTGA 549
QY 445 CFTCGAAAGTGGCGGAGAGACAGATATACTAAACACCTATCTCAAAAGAAAGAAAG 504
DB 550 AGTCGAAAGTGGCGGAGAGACAGATATACTAAACACCTATCTCAAAAGAAAGAAAG 609
QY 505 AAAAGAAATAGCTTATATATAGCTATATCTATGTTTATGTTTATGTTTATGTTTATAT 564
DB 610 AAAAGAAATAGCTTATATATAGCTATATCTATGTTTATGTTTATGTTTATGTTTATAT 668
QY 565 AAAGATCATCATATATGATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATATG 624
DB 669 AAAGATCATCATATATGATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATATG 728
QY 625 AGC 627
DB 729 AGC 731


```
QY 197 CAG-----GATCGTAGCCCTAGTCCA 220
Db 61 CAGGATCCGTAAGCCCTAGTACAGACCCGACAGACGTCGTACAGCCCTAGTCCA 120
QY 221 TATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAAC 280
Db 121 TATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAAC 180
QY 281 GAGTTTGAGAACCAACAAAGTGCATGTGCGAGCATTCGAACAGATCATGGAGAACCCAG 340
Db 181 GAGTTTGAGAACCAACAAAGTGCATGTGCGAGCATTCGAACAGATCATGGAGAACCCAG 240
QY 341 AGCGATAGTTGCGAGGGGCGCAACAGGACCAAGTTCAGAGGGAGCTCAGGAACCTTG 400
Db 241 AGCGATAGTTGCGAGGGGCGCAACAGGACCAAGTTCAGAGGGAGCTCAGGAACCTTG 300
QY 401 CCTCAACAGTGGCGCCCTTAGGGCAACACAGCGTTGCGACTTGCAGCTGCGAAAGTGGCGGC 460
Db 301 CCTCAACAGTGGCGCCCTTAGGGCAACACAGCGTTGCGACTTGCAGCTGCGAAAGTGGCGGC 360
QY 461 AGAGACAGATCTAAACACCTATCTCAAAAAAAGAAAGAAAGAAAGAAAGAAAGTCTTA 520
Db 361 AGAGACAGATCTAAACACCTATCTCAAAAAAAGAAAGAAAGAAAGAAAGAAAGTCTTA 420
QY 521 TATATAAGCTATTATCTATGTTGTTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAAT 580
Db 421 TATATAAGCTATTATCTATGTTGTTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAAT 479
QY 581 TGAATGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATAGGACCTTTAGAGTGC 640
Db 480 TGAATGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATAGGACCTTTAGAGTGC 539
QY 641 TTTT--ATGCGGTGTCTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAAT 698
Db 540 TTTTATGCGGTGTCTATGTTTGTGTCGAGAGTTGTAAACCATCTTGAATAAT 599
QY 699 AAAAAATCATGTTTGT 717
Db 600 NGAAGATCANGGTTTGT 618

RESULT 6
CD038306
LOCUS
DEFINITION
  UTPI005_D03 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION
  CD038306
VERSION
  CD038306.1 GI:30420144
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
  1 (bases 1 to 405)
  Location/Qualifiers
    seq primer: T3.
  AUTHORS
    Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
    and Lynch,R.E.
  TITLE
    Generation and Analyses of ESTs for Arachis hypogaea
  JOURNAL
    Unpublished (2003)
  COMMENT
    Contact: Baozhu Guo
    USDA/ARS, Crop Protection and Management Research Unit
    2747 Davis Rd., Tifton, GA 31794, USA
    Tel: 229-387-2334
    Fax: 229-387-2321
    Email: bguo@tifton.usda.gov
  FEATURES
    source
      1..405
        /organism="Arachis hypogaea"
        /mol_type="mRNA"
```

```
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPI005_D03"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). cDNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBlue-script phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."
```

ORIGIN

Query Match 43.9%; Score 315; DB 6; Length 405;
Best Local Similarity 99.7%; Pred. No. 1.4e-67;
Matches 326; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 391 CAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCAACCAAGCTTGGAGCTTGGACCTCGA 450
Db 77 CAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCAACCAAGCTTGGAGCTTGGACCTCGA 136
QY 451 AAGTGGCGCGAGAGACAGATCTAAACACCTATCTCAAAAAAAGAAAGAAAGAAAGA 510
Db 137 AAGTGGCGCGAGAGACAGATCTAAACACCTATCTCAAAAAAAGAAAGAAAGAAAGA 196
QY 511 AATAGCTTATATATAAGCTATTATCTATGTTTGTGTTAGTTTGGTATATATAGAT 570
Db 197 AATAGCTTATATATAAGCTATTATCTATGTTTGTGTTAGTTTGGTATATATAGAT 255
QY 571 CATCATATATCAATGCTGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATGACACC 630
Db 256 CATCATATATCAATGCTGTTGATCGTGTAACTAAGGCAAGCTTAGTGTATATGACACC 315
QY 631 TTTAGAGTCTTTTATGCGCTTGTCTATGTTTGTGCTGCAGAGTTGTAACCATCTTGA 690
Db 316 TTTAGAGTCTTTTATGCGCTTGTCTATGTTTGTGCTGCAGAGTTGTAACCATCTTGA 375
QY 691 AATAATATAAAGATCATGTTTGT 717
Db 376 AATAATATAAAGATCATGTTTGT 402

RESULT 7
CD038459
LOCUS
DEFINITION
 UTPI007_F09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
ACCESSION
 CD038459
VERSION
 CD038459.1 GI:30420297
KEYWORDS
 EST.
SOURCE
 Arachis hypogaea (peanut)
 Arachis hypogaea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
 Aeschynomeneae; Arachis.
 1 (bases 1 to 308)

AUTHORS Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
TITLE Generation and Analyses of ESTs for *Arachis hypogaea*
JOURNAL Unpublished (2003)
COMMENT Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
source
1..308
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTPI007_F09"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="Xl1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar Al3 (NCV11XR4). Al3 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIzol-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match 40.3%; Score 289; DB 6; Length 308;
Best Local Similarity 99.7%; Pred. No. 4.1e-61;
Matches 300; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 417 TTAGGGCACCACAGCGTTGCGACTTGGACGTCGAAAGTCGAAAGTCGCGGACAGACAGATCTAA 476
Db 1 TTAGGGCACCACAGCGTTGCGACTTGGACGTCGAAAGTCGCGGACAGACAGATCTAA 60
QY 477 CACCTATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 536
Db 61 CACCTATCTCAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 120
QY 537 TATGGTTATGTTAGTTTGGTAAATAATAAGATCATCACTATATGAATGTGTGATCGT 596
Db 121 TATGGTTATG- TTAGTTTGGTAAATAATAAGATCATCACTATATGAATGTGTGATCGT 179
QY 597 GTTAACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGTCT 656
Db 180 GTTAACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGTCT 239
QY 657 ATGTTTGTGCTGAGAGTTGTAACCATCTTGAATAATAATAATAATAATAATAATAATAATAATAATA 716
Db 240 ATGTTTGTGCTGAGAGTTGTAACCATCTTGAATAATAATAATAATAATAATAATAATAATAATAATA 299
QY 717 T 717
Db 300 T 300

RESULT 8
C0897504
LOCUS
DEFINITION
C0897504 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.
ACCESSION
C0897504
VERSION
C0897504.1 GI:51237294
KEYWORDS
EST.
SOURCE
Arachis hypogaea (peanut)
ORGANISM
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 375)
Yan,Y.S., Wang,L. and Huang,S.Z.
Expressed sequence tags of mid-matured cotyledons of peanut
Unpublished (2004)
Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (86)02084036592
Email: yy9803@yahoo.com.cn
Seq primer: pTriplEx2 Forward.
Location/Qualifiers
source
1..375
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/tissue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/note="Organ: Seed; Vector: lambdaTriplEx2"

ORIGIN
Query Match 39.5%; Score 283.4; DB 7; Length 375;
Best Local Similarity 88.8%; Pred. No. 1e-59;
Matches 333; Conservative 0; Mismatches 6; Indels 36; Gaps 1;
QY 103 CGAGAGGGCGAACCCTGAGGCCCTGCGAGCAACATCTCATGAGAGATCCAAACGTGACGA 162
Db 1 CGAGAGGGCGAACCCTTAGGCCCTGCGAGCAACATCTCATGAGAGATCCAAACGTGACGA 60
QY 163 GGATTTCATATGAA CGGAGCCCGTACAGCCCTAGTCAG----- 199
Db 61 GGATTTCATATGAGCGGAGCCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCAGGA 120
QY 200 -----GATCGGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGAATCCTC 246
Db 121 CCCGACAGACAGTGTATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGAATCCTC 180
QY 247 TCAGACCAAGAGAGGTTTGCATGAGCTGAACAGAGTTTGAGAAACAACCAAGGTGCAT 306
Db 181 TCAGACCAAGAGAGGTTTGCATGAGCTGAACAGAGTTTGAGAAACAACCAAGGTGCAT 240
QY 307 GTGCGAGGCTTGCACAGATCATGGAGAACACAGAGCGATAGTTGCGAGGGAGGCAACA 366
Db 241 GTGCGAGGCTTGCACAGATCATGGAGAACACAGAGCGATAGTTGCGAGGGAGGCAACA 300
QY 367 GGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTCGCGCTTTAGGGCACC 426
Db 301 GGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTCGCGCTTTAGGGCACC 360
QY 427 ACAGCGTTGCGACTT 441
Db 361 ACAGCGTTGCGACTT 375

RESULT 9

```

CD038415      696 bp      mRNA      linear      EST 07-MAY-2003
LOCUS      UTPPI006_H11 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION      Arachis hypogaea cDNA clone UTPPI006_H11 5', mRNA sequence.
ACCESSION      CD038415
VERSION      CD038415.1 GI:30420253
KEYWORDS      EST.
SOURCE      Arachis hypogaea (peanut)
ORGANISM      Arachis hypogaea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE      1 (bases 1 to 696)
AUTHORS      Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
              and Lynch,R.E.
TITLE      Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
              Location/Qualifiers
                1..696
                  /organism="Arachis hypogaea"
                  /mol_type="mRNA"
                  /cultivar="A13"
                  /db_xref="taxon:3818"
                  /clone="UTPP1006_H11"
                  /tissue_type="Immature pods"
                  /dev_stage="R6"
                  /lab_host="XLI-blue"
                  /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
                  (UTPP)"
                  /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
                  cDNA library was constructed from peanut cultivar A13
                  (NCV11XAR4). A13 has resistance to Aspergillus infection
                  and drought tolerance. The immature pods that developed to
                  R6 stage were collected from different plants, and placed
                  into liquid N2 immediately and stored in -80oC freezer.
                  Total RNA was isolated with TRIzol-Reagent
                  ultrapure (GIBCOBRL). mRNA was extracted and purified from
                  total RNA (Promega). cDNA synthesis and library
                  construction followed the protocol of by ZAP-cDNA Gigapack
                  III Gold cloning kit (Stratagene). The cDNA above 500bp
                  were collected after size-fraction. The inserts were
                  directionally cloned into Uni-ZAP XR vector using XhoI
                  EcoRI sites adapters. The lambda library was packed into
                  phages using Gigapack III Gold (Stratagene). The
                  un-amplified library was used to excise pBluescript
                  phagemids from the Uni-ZAP XR vector, and the phagemids
                  was used to transform the host bacteria SOLR. The library
                  was constructed by Dr. Meng Luo and Dr. Phat Dang."

FEATURES             source
  source
    1..732
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPP1006_F02"
      /tissue_type="Immature pods"
      /dev_stage="R6"
      /lab_host="XLI-blue"
      /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
      (UTPP)"
      /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
      cDNA library was constructed from peanut cultivar A13
      (NCV11XAR4). A13 has resistance to Aspergillus infection
      and drought tolerance. The immature pods that developed to
      R6 stage were collected from different plants, and placed
      into liquid N2 immediately and stored in -80oC freezer.
      Total RNA was isolated with TRIzol-Reagent

ORIGIN
  Query Match      31.2%; Score 223.8; DB 6; Length 696;
  Best Local Similarity 71.9%; Pred. No. 8.1e-45;
  Matches 346; Conservative 0; Mismatches 107; Indels 28; Gaps 3;

QY 219 CATATGATCGGAGGCGCTGCGATCCTCTCAGCACCAAGAGAGGCTGTTCGAATGAGCTGA 278
DB 234 CCTACCATATTAGAGTACTCGATCCTCCGACCAGCACAGAGGCTCTCGGATGAGCTGA 293
QY 279 ACAGTTTGAACAACACCAAGGTGATGTGCGAGGCGATTGCAACAGATCATGGAGAACC 338
DB 294 ACCAGATGAGAAACACACAGAGATGCGATGTGCGAGGCGATTGCGAGCAGATAATGAGAACC 353
QY 339 AGAGCCATAGTTGCGAGGAGGCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCT 398

```

```

DB 354 AGTGGCATAGGTTGCAGGACAGGCAAAATGGTCAGCAGTTTCAAGAGAGAGCTCATGAACCT 413
QY 399 TGCCTCAACAGTGGCGGCTTAGGGACACACAGCGTTGGACGTCGAAAGTGGCG 458
DB 414 TGCCCCAACAGTGTAACTTTTAGGGACACACAGCGTTGGCATTTGGACGT-----GA 464
QY 459 GCAGAGACAGACTAAACACACCTATCTCAAAAAGAGAAAAGAAAAGAAAAGAAATAGCT 518
DB 465 GTGGCGGCAGATGCTAGACTCAAAAATAATAATCTGTGCCAAAACAACACTAGTAGGAAGT 524
QY 519 TATATATAAGCTATTATCTATGTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT 578
DB 525 AGCTTATGAGCTATTATGTTATGTTT-----GTTTCGTTTAAATAAACAATCATCACTG 577
QY 579 TATGAATGTTGATCGTGTAACTAAAGCAAGCTTAGGTTATATAGACACCTTTTAGT 638
DB 578 TATGAATGTTGTCAT-----AGCTAGGTAAGGTTATATGAGCCCTTCGGTGT 625
QY 639 GCTTTTATGGCGTGTCTATGTTTGTCTGCAGAGTTGTAACCATCTTGAATAATAT 698
DB 626 GTCTTATGCGCTTTACCTGTGTTTGTCTACTGCAAAAGTTTAAACCAACCAATGAATAAAG 685
QY 699 A 699
DB 686 A 686

RESULT 10
CD038394
LOCUS      UTPPI006_F02 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
DEFINITION      Arachis hypogaea cDNA clone UTPPI006_F02 5', mRNA sequence.
ACCESSION      CD038394
VERSION      CD038394.1 GI:30420232
KEYWORDS      EST.
SOURCE      Arachis hypogaea (peanut)
ORGANISM      Arachis hypogaea
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
              Aeschynomeneae; Arachis.
REFERENCE      1 (bases 1 to 732)
AUTHORS      Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
              and Lynch,R.E.
TITLE      Generation and Analyses of ESTs for Arachis hypogaea
JOURNAL      Unpublished (2003)
COMMENT      Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.
              Location/Qualifiers
                1..732
                  /organism="Arachis hypogaea"
                  /mol_type="mRNA"
                  /cultivar="A13"
                  /db_xref="taxon:3818"
                  /clone="UTPP1006_F02"
                  /tissue_type="Immature pods"
                  /dev_stage="R6"
                  /lab_host="XLI-blue"
                  /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
                  (UTPP)"
                  /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
                  cDNA library was constructed from peanut cultivar A13
                  (NCV11XAR4). A13 has resistance to Aspergillus infection
                  and drought tolerance. The immature pods that developed to
                  R6 stage were collected from different plants, and placed
                  into liquid N2 immediately and stored in -80oC freezer.
                  Total RNA was isolated with TRIzol-Reagent

```

ultrapure(GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lamda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 30.0%; Score 215.2; DB 6; Length 732;
Best Local Similarity 72.0%; Pred. No. 1.1e-42;
Matches 347; Conservative 0; Mismatches 108; Indels 27; Gaps 4;
QY 219 CATATGATCGGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTGA 278
DB 262 CCTACGATATTAGGAGTACTCGATCCTCCAGCAGCAAGAGGTGCTCGATGAGCTGA 321
QY 279 ACAGTTTCAGAACCAACCAAGGTGATGTCGAGGCATTCGACAGATCATGGAGAACC 338
DB 322 ACAGATGGAGAACACACAGAGATGCATGTGCGAGGCATTGCGAGCAGATAATGGAGAACC 381
QY 339 AGAGCGCATAGTTGTCAGGGGAGGCAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAAC 398
DB 382 AGTGCATAGTTGTCAGGACAGCAAAATGTCGAGCATTCGAGAGAGCTCATGAAC 441
QY 399 TGCTCAACAGTGCAGCGCTTAGGGCACCAAGCGTTGCGACTTTGGACGTTCGAAAGTGGG 458
DB 442 TGCCCCAACAGTGAATCTTTAGGGCACCAAGCGTTGCGACTTTGGACGTTCGAAAGTGGG 498
QY 459 GCAGAGACAGATCACTAAACACCTATCTCAAAAAGAAAGAAAGAAAGAAATAGCT 518
DB 499 GCAGATGCTAGACTCAAAAATATATATCTGTGCCAAAACAAACTTAGTAGGAAGTAGCT 558
QY 519 TATATATAGCTATTATCTATGTTATGTTTATGTTTGTGTTATATATATATATATATATAT 578
DB 559 ATATTATAGCTATTATGTTATGCTT-----GTTTCGTTAATATATATATATATATAT 611
QY 579 TATGAATGTTGATCGTGTAACTAAGCAAGCTTAGGTTATATAGCACCTTTTAGAGT 638
DB 612 TATGAATGTTGAT-----AGGTAAGGTTATATAGCACCTTCGGTGT 655
QY 639 GCTTTTATGCG-GTTGCTCTATGTTTGTGTCGAGAGTTGTAACCATTTGAAATATA 697
DB 656 GCTCTATGGCTTTTACCTATGTTTGTCTACTGCAAAAGTTTTAACCATGAAATATAA 715
QY 698 TA 699
DB 716 GA 717

RESULT 11
CD038326

LOCUS CD038326 661 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPI005 F10 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPI005_F10 5', mRNA sequence.

ACCESSION
CD038326

VERSION CD038326.1 GI:30420164

KEYWORDS
SOURCE

Arachis hypogaea (peanut)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 661)

REFERENCE
AUTHORS

Luo, M., Dang, P., Guo, B. Z., Holbrook, C. C., Lee, R. D., Bausher, M. G.
and Lynch, R. E.

TITLE
Generation and Analyses of ESTs for Arachis hypogaea

JOURNAL
COMMENT

Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: n3

FEATURES
source

1..661
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Al3"
/db_xref="taxon:3818"
/clone="UTPI005_F10"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
cDNA library was constructed from peanut cultivar Al3
(NCV11XAR4). Al3 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIZOL-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lamda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise pBluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 29.1%; Score 208.8; DB 6; Length 661;
Best Local Similarity 65.1%; Pred. No. 4.4e-41;
Matches 457; Conservative 0; Mismatches 164; Indels 81; Gaps 7;
QY 4 CACCATATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACCGCATCTGC---GAGGCA 60
DB 25 CACCATCTGTGTAGCCCTCTTGGCCCTGCTGCTGGTCACACGCTCCCGGATGAGCGG 84
QY 61 GCAGTGGGAACTCCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACT 117
DB 85 CGAGAGGGGGGACAAAGGGGACTCATCAAGCTGCGAGAGGCGAGGTAGACAGGTTGAAC 144
QY 118 GAGGCCCTGCGAGCAACATCTCATGCAAGAATCCAACTGACGAGGATTCATATGAACG 177
DB 145 CAAGCCCTGCGAGCAGCACATAATGCAGAGGATAATGGCGAGCAAGCAGTAGTACGACT- 203
QY 178 GGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGGCGC 237
DB 204 -----CCTACGATATTAGAGTAC 222
QY 238 TGGATCTCTCAGCACCAAGAGAGGTTGTGCAATGAGCTGAACGATTTAGAACCAACA 297
DB 223 TCGATCTTCCAGCCAGCAACAGAGGTGCTGCGATGAGCTGGACCCAGATGGAGNACACAGA 282
QY 298 AAGGTGCAATGTGCGAGGCATTTGCAACAGATCATGAGAACACAGACCGATAGTTCCAGGG 357
DB 283 GAGATGCAATGTGCGAGGCATTTGCGAGCAGATAATGAGAAACAGTGCATAGTTGCAGGA 342
QY 358 GAGGCAACAGAGNACACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGCT 417
DB 343 CAGGCAAAATGGTGCAGCAGTTCAAGAGAGAGCTCATGAACCTTGCTCAACAGTGTAACTT 402


```

VERSION      CD038092.1  GI:30419930
KEYWORDS     EST.
SOURCE       Arachis hypogaea (peanut)
ORGANISM     Arachis hypogaea

REFERENCE    Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
              and Lynch, R.E.
              Generation and Analyses of ESTs for Arachis hypogaea
              Unpublished (2003)
              Contact: Baozhu Guo
              Molecular Genetics
              USDA/ARS, Crop Protection and Management Research Unit
              2747 Davis Rd., Tifton, GA 31794, USA
              Tel: 229-387-2334
              Fax: 229-387-2321
              Email: bguo@tifton.usda.gov
              Seq primer: T3.

FEATURES     Location/Qualifiers
              1..680
                 /organism="Arachis hypogaea"
                 /mol_type="mRNA"
                 /cultivar="A13"
                 /db_xref="taxon:3818"
                 /clone="UTPPI002.D04"
                 /tissue_type="Immature pods"
                 /dev_stage="R6"
                 /lab_host="XLI-blue"
                 /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
                 (UTPP)"
                 /note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
                 cDNA library was constructed from peanut cultivar A13;
                 (NCV1XA4). A13 has resistance to Aspergillus infection
                 and drought tolerance. The immature pods that developed to
                 R6 stage were collected from different plants, and placed
                 into liquid N2 immediately and stored in -80oC freezer.
                 Total RNA was isolated with TRIzol-Reagent
                 ultrapur (GIBCOBRL). mRNA was extracted and purified from
                 total RNA (Promega). cDNA synthesis and library
                 construction followed the protocol of by ZAP-cDNA Gigapack
                 III Gold cloning kit (Stratagene). The cDNA above 500bp
                 were collected after size-fraction. The inserts were
                 directionally cloned into Uni-ZAP XR vector using XhoI
                 EcoRI sites adapters. The lambda library was packed into
                 phages using Gigapack III Gold (Stratagene). The
                 un-amplified library was used to excise phagescript
                 phagemids from the Uni-ZAP XR vector, and the phagemids
                 was used to transform the host bacteria SOLR. The library
                 was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match      18.6%; Score 133.2; DB 6; Length 680;
Best Local Similarity 58.3%; Pred. No. 3e-22;
Matches 344; Conservative 0; Mismatches 213; Indels 33; Gaps 5;

QY 76 AGAGACAGAAGATGCCAGGACGAGCTCGAGAGGGGGAACCTGAGGCCCTGGAGCAACA 135
DB 107 AGGGGATACCACTAGTCAGAGGCAGTTGACAGAGGGGCAACCTGAGGCCCTGTGAGGAACA 166
QY 136 TCTCATGCAAGATGCCAAGCTGACGAGGATTCATATGAACGGACCCGTCACAGCCCTAG 195
DB 167 CATAGGCAAGGGTGGACCAAGACGACGACGACGACGATACCCCTACAGCCA 226
QY 196 TCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCCTCTCAGCACC 255
DB 227 ACGGGATCCAGAGGACGACGACCC-----GGCGAATCTGACGAAGATCAAGA 274
QY 256 AGAGAGGTGTTTCGATGAGCTGAACGAGTTTGACACACCAACGATGATGTCGAGGC 315
DB 275 GCAAGAGGTGCTGCAACGAGCTCAACCGGTTCCAGGAATAACCAAGGTGCATGTGCCAGGC 334

```

```

QY 316 ATTGCAACAGATCATGGAGAACCCAGAGCGATAGGTT---GCAGGGAGGCAACAGGAGCA 372
DB 335 ACTTCAAAGATCTCCAGAACCCAGAGCTTTAGGTTCCAGCAGGACAGGACCGAGTTGCA 394
QY 373 ACAGTTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTGGCGGCTTAGGGCACCACAGCG 432
DB 395 TCAGATGGAGGGAGCTCAGGAATTCGCCCATAACTCGCGGTTTCAGGTCAACCAAGCG 454
QY 433 TTGCGACATT-----GGACGTGAAAGTGGCGCAGAGACAGATACATAACACCC 480
DB 455 TTGCGACCTTAGTAGCCGACGCCCTACTAAACAGACTAGCATTGTCGTTTAAATTGCG 514
QY 481 TATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGCTATTATCTATG 540
DB 515 TTACCCACACAGAAATAACAATGATATAATAATAGCTTTTTTACAGCTATTCTA 574
QY 541 GTTATGTTTGGTAAATAAAGATCATCATCTATAGA--ATGTTTGTATCGTGT 598
DB 575 TGTCTATGTTGTTTGGTAGCAATAAAGATCATCACCATTATGATAGTGTGATCGTAT 634
QY 599 TAACTAAGCAAGCTTAGGTTATATAGCACCTTTAGAGTGTCTTTATGG 648
DB 635 TGTCTGTGCGGAAGTTATAT-----GGGCACCTTTAAATGTGCTTTTATGG 680

RESULT 15
CO897507
LOCUS     EST000012 Peanut Lambda Express library Arachis hypogaea cDNA 5',
DEFINITION mRNA sequence.
ACCESSION CO897507
VERSION    CO897507.1 GI:51237297
KEYWORDS   EST.
SOURCE     Arachis hypogaea (peanut)
ORGANISM   Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE  1 (bases 1 to 351)
AUTHORS   Yan, Y.S., Wang, L. and Huang, S.Z.
TITLE      Expressed sequence tags of mid-matured cotyledons of peanut
JOURNAL    Unpublished (2004)
COMMENT    Contact: Shangzhi Huang
            Plant Development and Molecular Biology
            Sun Yat-sen University, Department of Biology and Biotechnology
            Guangzhou, 510275, China
            Tel: (860)02084036592
            Email: YYS9803@yahoo.com.cn
            Seq primer: pTriplex2 Forward.
FEATURES   Location/Qualifiers
            1..351
               /organism="Arachis hypogaea"
               /mol_type="mRNA"
               /cultivar="Yueyou 523"
               /db_xref="taxon:3818"
               /tissue_type="Cotyledons"
               /dev_stage="Mid-matured stage"
               /lab_host="E.coli BM25.8"
               /clone_lib="Peanut Lambda Express library"
               /note="Organ: Seed; Vector: lambdaTriplex2"

ORIGIN
Query Match      18.4%; Score 131.8; DB 7; Length 351;
Best Local Similarity 80.6%; Pred. No. 6.1e-22;
Matches 154; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 219 CATATGATCGGAGAGCGCTCGATCCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGA 278
DB 161 CCTAGATATTAGGAGTACTCGATCCTCCGACCACACAGAGGTGCTCGATGAGCTGG 220
QY 279 ACGAGTTTGAGAACCAACCAAGGTGTCATGTGCGAGGCATTGCAACAGATCATGGAGAAC 338

```

[illegible]

Search completed: August 24, 2005, 03:44:12
Job time : 1685.67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 117.962 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gtcaccatactagtagcc.....taaaagatcatgtttgttt 717

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	100.0	717	4	US-09-106-872A-20
2	717	100.0	717	4	US-09-191-593-6
3	717	100.0	717	4	US-09-191-593-9
4	713.8	99.6	743	4	US-09-106-872A-1
5	65.6	9.1	777	2	US-08-618-911-3
6	64	8.9	777	2	US-08-618-911-5
7	63.4	8.8	770	3	US-08-938-675A-1
8	63.4	8.8	770	4	US-09-531-727-1
9	58.8	8.2	7218	1	US-08-232-463-14
10	54.4	7.6	723	2	US-08-618-911-1
11	44.4	6.2	168104	4	US-09-949-016-12026
12	44.4	6.2	168105	4	US-09-949-016-16554
13	44.4	6.2	1664976	4	US-08-916-421B-1
14	44.4	6.2	1664976	4	US-09-692-570-1
15	44.2	6.2	56685	4	US-09-949-016-14026
16	44	6.1	31739	4	US-09-949-016-16226
17	43.6	6.1	601	4	US-09-949-016-44480
18	43.6	6.1	34855	4	US-09-949-016-13004
19	43.6	6.1	56302	4	US-09-949-016-11892
20	43.2	6.0	251672	4	US-09-949-016-17296
21	43.2	6.0	251682	4	US-09-949-016-11973
22	42.6	5.9	133613	4	US-09-949-016-15824
23	42.4	5.9	289	3	US-09-007-005-17
24	42.4	5.9	289	3	US-09-244-796-3
25	42.2	5.9	95890	4	US-09-949-016-16412
26	42	5.9	5630	2	US-08-937-931-1
27	42	5.9	5630	3	US-09-285-502-1

28 5.9 5630 3 US-09-709-126-1 Sequence 1, Appli
29 5.9 5630 3 US-09-871-385A-1 Sequence 1, Appli
30 5.9 87752 4 US-09-949-016-16807 Sequence 16807, A
31 42 5.9 197875 4 US-09-949-016-15425 Sequence 15425, A
32 41.6 5.8 601 4 US-09-949-016-88013 Sequence 88013, A
33 41.6 5.8 717 4 US-09-106-872A-20 Sequence 20, Appli
34 41.6 5.8 717 4 US-09-191-593-6 Sequence 6, Appli
35 41.6 5.8 717 4 US-09-191-593-9 Sequence 9, Appli
36 41.6 5.8 743 4 US-09-106-872A-1 Sequence 1, Appli
37 41.6 5.8 62873 4 US-09-949-016-15676 Sequence 15676, A
38 41.6 5.8 64610 4 US-09-949-016-12214 Sequence 12214, A
39 41.6 5.8 72128 4 US-09-949-016-16018 Sequence 16018, A
40 41.6 5.8 94879 4 US-09-949-016-12101 Sequence 12101, A
41 41.6 5.8 94884 4 US-09-949-016-13393 Sequence 13393, A
42 41.6 5.8 670689 4 US-09-949-016-12505 Sequence 12505, A
43 41.6 5.8 670690 4 US-09-949-016-14207 Sequence 14207, A
44 41.4 5.8 30569 4 US-09-949-016-12592 Sequence 12592, A
45 41.2 5.7 601 4 US-09-949-016-88012 Sequence 88012, A

ALIGNMENTS

RESULT 1

US-09-106-872A-20
; Sequence 20, Application US/09106872A

; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannion, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Atachis hypogaea
US-09-106-872A-20

Query Match 100.0%; Score 717; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.9e-189;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCGCGCATCTCGGAGGCA 60
Db 1 GCTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCGCGCATCTCGGAGGCA 60
Qy 61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAAGCTGAG 120
Db 61 GCAGTGGGAATCTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAAGCTGAG 120
Qy 121 GCCTCGCAGCAACATCTCATGAGAAATCCAACTGACGAGGATTCATATGAAACGGGA 180
Db 121 GCCTCGCAGCAACATCTCATGAGAAATCCAACTGACGAGGATTCATATGAAACGGGA 180
Qy 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCCCTGG 240
Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGGCCCTGG 240
Qy 241 ATCCTCTCAGCACCAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAG 300
Db 241 ATCCTCTCAGCACCAGAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAG 300

QY 301 GTGCATGTGGAGGCAATTGCAACAGATCATGGAGAACGAGCGATAGGTTGCAGGGAG 360
Db 301 GTGCATGTGGAGGCAATTGCAACAGATCATGGAGAACGAGCGATAGGTTGCAGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGGAGGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGGAGGCTCAGAACTTGCCTCAACAGTGGCGCTTAG 420
QY 421 GGCAACACAGCGTTGGAGCTTGGACGTCGAAAGTGGCGGAGAGAGAGATCACTAAACACC 480
Db 421 GGCAACACAGCGTTGGAGCTTGGACGTCGAAAGTGGCGGAGAGAGAGATCACTAAACACC 480
QY 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
QY 541 GTTATGTTAGTTTGGTAAATAATAAAGATCATCATATGATGATGATGATGATGATGATGATG 600
Db 541 GTTATGTTAGTTTGGTAAATAATAAAGATCATCATATGATGATGATGATGATGATGATGATG 600
QY 601 ACTAAGGCAAGCTTAGCTTATATAGCACCCTTTAGAGTGCTTTTATGGCGTGTCTATGT 660
Db 601 ACTAAGGCAAGCTTAGCTTATATAGCACCCTTTAGAGTGCTTTTATGGCGTGTCTATGT 660
QY 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAATAATAATAATAATAATAATAATAATA 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAATAATAATAATAATAATAATAATAATAATA 717

RESULT 2

US-09-191-593-6
; Sequence 6, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 717 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; DESCRIPTION: Identified as Ara h II P38
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: No. 6835824 applicable
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE: Clone P38
; DEVELOPMENTAL STAGE: Seed
; HAPLOTYPE: No. 6835824 applicable
; TISSUE TYPE: Seed mRNA, cDNA library
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLE: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY: Florunner seed cDNA expression
; LIBRARY: library in Uni-ZAP XR vector
; CLONE: P38
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..475
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-6
Query Match 100.0%; Score 717; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.9e-189;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTCACCATACTAGTAGCCCTGCGCCCTTTTCCTCTCGCTGCCCAAGGAGGCGGAACTCGAG 60
Db 1 GCTCACCATACTAGTAGCCCTGCGCCCTTTTCCTCTCGCTGCCCAAGGAGGCGGAACTCGAG 60
QY 61 GCAGTGGGAACCTCCAAAGGAGAGAGAGATGCCAGAGCAGAGTCCAGAGAGGCGGAACTCGAG 120
Db 61 GCAGTGGGAACCTCCAAAGGAGAGAGAGATGCCAGAGCAGAGTCCAGAGAGGCGGAACTCGAG 120
QY 121 GCCCTGCGAGCAACAATCTCATGCAAGATCCAAAGTCCAGAGGATTCATATGAACGGGA 180
Db 121 GCCCTGCGAGCAACAATCTCATGCAAGATCCAAAGTCCAGAGGATTCATATGAACGGGA 180
QY 181 CCCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTAGGATCCGT 240
Db 181 CCCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTAGGATCCGT 240

QY 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAACCAACCAAG 300
DB 241 ATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAACGAGTTTGAGAACCAACCAAG 300
QY 301 GTGATGTGCGAGGCAATTCACACAGATCATGGAGAACCAAGAGCGATAGTTTGACAGGGGAG 360
DB 301 GTGATGTGCGAGGCAATTCACACAGATCATGGAGAACCAAGAGCGATAGTTTGACAGGGGAG 360
QY 361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTAG 420
DB 361 GCAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTAG 420
QY 421 GGCAACACAGCTTGCACCTTGGACCTGGAAGTGGCGGCGAGACAGATACCTAAACACC 480
DB 421 GGCAACACAGCTTGCACCTTGGACCTGGAAGTGGCGGCGAGACAGATACCTAAACACC 480
QY 481 TATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 540
DB 481 TATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 540
QY 541 GTTATGTTTGTAGTTTGTATATAATAAAGATCATCATATATATATATATATATATATATAT 600
DB 541 GTTATGTTTGTAGTTTGTATATAATAAAGATCATCATATATATATATATATATATATATAT 600
QY 601 ACTAAGCGACGCTTAGCTTATATAGCACCTTTAGAGTGTCTTTATGCGCTGCTCTATGT 660
DB 601 ACTAAGCGACGCTTAGCTTATATAGCACCTTTAGAGTGTCTTTATGCGCTGCTCTATGT 660
QY 661 TTTGTTCTGCGAGAGTTGTAACCACTTTGAAATATATAAAGATCATCATGTTTGT 717
DB 661 TTTGTTCTGCGAGAGTTGTAACCACTTTGAAATATATAAAGATCATCATGTTTGT 717

RESULT 3

US-09-191-593-9
; Sequence 9, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996

ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 717 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
DESCRIPTION: identified as Ara h II cDNA clone
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: No. 6835824 applicable
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE: seed
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE: seed cDNA
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY: florunner seed cDNA expression
LIBRARY: library in Uni-ZAP XR vector
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
OTHER INFORMATION: allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-9

Query Match 100.0%; Score 717; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 2,9e-189;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGTGGCCACGCGCATCTCGAGGCA 60
DB 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGTGGCCACGCGCATCTCGAGGCA 60
QY 61 GCAGTGGGAACTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
DB 61 GCAGTGGGAACTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAG 120
QY 121 GCCTCGGAGGACACATCTCATGCAGAGATCCAACTGACGAGATTCATATGAACGGGA 180


```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B98-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
;
US-09-531-727-1

Query Match      8.8%; Score 63.4; DB 4; Length 770;
Best Local Similarity 51.6%; Pred. No. 1.4e-07;
Matches 369; Conservative 0; Mismatches 301; Indels 45; Gaps 8;

QY      7  CATACTAGTAGCCCTCGCCCTTTTCTCTCTCGCTGCCCGCCAGCATCTCGGAGGCAGCATG 66
DB      28  CACAACTCTCTCATCTCTCTTCTCTTCTGTGATGCCCCACACTTGCAGCGCTTCAAAATG 87

QY      67  GGAACTCCAAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGGCCCTG 126
DB      88  GCAGCACCCAGCAAGA---TAGTGC CGCAGCAGCTCCAGGGGTGAACCTCAGCCCTG 144

QY      127  CGAGCAACATCTCATGCAAGAAGATCCAACTGTGACGAGGATTCATATGAA CGGGACCCGTA 186
DB      145  CGAAGAAGCATCATGTGAGAAGATCCAAAGGCCGCGCGATGACGATGATGATGATGATGACGA 204

QY      187  CAGC-----CCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCCGAG----- 231
DB      205  CGACAATCAATCTTCAGGACCATGCGGGAGAGATCACTACATAAGGAGGAACGAAGG 264

QY      232  -----AGCGCTGGATCTCTCAGCACCAAGAGAGAGGTGTGTCATAGCTGAAACGA 282
DB      265  AAAAAGACGAAGACGAAGAAGAAGAGACACATGCAGAAGTGTCTGCACAGAAATGAGCGA 324

QY      283  GTTTTGAGAACAAACCAAGGTGCATGTGCGAGGCATTCCAACAGATCATGAGAACACGAG 342
DB      325  GCT---GAGAAAGCCCAATGCCNGTGCAAGCGCTGCAAGAGATATGGAGAACGAG 381

QY      343  CGATAGGTTTCGAGGGGAGCAACAGGAGCAACAGTTCGAAGAGGAGCTCAGGAACCTTGCC 402
DB      382  CGAGGAACTCGAGGAGAAGCAGAAGAAGAAA---ATGGAAGAGGAGCTCACTTAACCTTGGC 438

QY      403  TCAACAGTGGCGCTTAGGGCACCAACAGCGTTCGCACTTGGACGTGCA--AGTGGCGGCA 461
DB      439  TACTATGTGCAGGTTTGGACCCATGATCCAGTGGCAGCTTGTCTCCGATGACATAAGAAAGT 498

QY      462  GAGACAGATACTAAACACCTATCTCAAAAAAGAAAAAGAAAAAGAAAAAGAAATAGCTTAT 521
DB      499  TAAAGCAATGTTGTCACTTGTGCTACTAACAATGATGTGATAGTTATTCCTAGCT--- 555

QY      522  ATATAAGCTATTATCTATGGTTATGTTTAGTTTGGTAAATAATAAGATCATCATATAT 581
DB      556  -----AGCTATAACATAAGCTGTCTCTGAGTGTGTGTGTATATTAATAAAGATCATCATCTG 610

QY      582  GAATGTGTTCAATCGTGTTAACTAAGGCAAGCTTAGGTTATATATGAGCACCTTTAGAGTGCT 641
DB      611  GTGAATGGTGATCGTGTACGTACCTACTTAGTAGGCAATGGAAGCACTTAGAGTGTGCT 670

QY      642  TT---TATGGCGTTGCTATGTTTGTGTCGAGAGTTGTAACCATCTTTGAAAT 693
DB      671  TTGTGATGCGCTTGCCTCTGTTTGTGAGACTTTTGTAAATGTTTTCGAGTTTAAAT 725

```

[illegible]

```
QY 489 AAAAAAGAAAAAGAAAAAGAAA 512
Db 1089 RRRRRRRRRRRRRRRRRRA 1066

RESULT 10
US-08-618-911-1
; Sequence 1, Application US/08618911
; Patent No. 5850016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; TITLE OF INVENTION: SEEDS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..474
US-08-618-911-1

Query Match 7.6%; Score 54.4; DB 2; Length 723;
Best Local Similarity 61.8%; Pred. No. 4.3e-05;
Matches 123; Conservative 0; Mismatches 71; Indels 6; Gaps 2;

QY 251 CACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTTGAGAACACCAAGGTCATGTGC 310
Db 277 CACATCGAGAGTGTTCGACGGAATGAGGAGCT---GAAAGCCCCATATGCCAGTGC 333

QY 311 GAGGCATTGCAACAGATCATGTGGAGAACACGAGCGATAGGTTCAGGGGAGGCAACAGGAG 370
Db 334 AAAGCGCTACAGAAGATAATGGTATACCGAGCGGAGCAACTGGAGGGGAAGAGAGAAG 393

QY 371 CAACAGTTCAAGAGGAGCTCAGGAACCTTCACAGTGGCGCCCTTAGGGCACCACAG 430
Db 394 ---CAGATGGAGAGAGAGCTCATGAACCTTGGCTATTAGGTGCAGGTGGGACCCATGATA 450

QY 431 CGTTGCGACTTGGACGTCGA 450
Db 451 GGGTGGCACTTGTCCTCCGA 470

us-09-949-016-12026/c
; Sequence 12026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12026
; LENGTH: 168104
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168104)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12026

Query Match 6.2%; Score 44.4; DB 4; Length 168104;
Best Local Similarity 59.5%; Pred. No. 0.35;
Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 386 GAGCTCAGGAACCTGCCTCAACAGTGGCGCCTTAGGGCACCACAGGTTGCCGACTTGGAC 445
Db 120712 GAGCCAGGAAGTCAAGGCTGCAGTGAGCCATGATCACCACCTGCAITTCAGCCTGGGT 120653

QY 446 GTCGAAAGTGGCGGAGACAGACAGATACATACTAAACACCTATCTCAAAAAAGAAAAAGAAA 505
Db 120652 GACAGATTAAGACCCTGTCTCAAAAAACAAAAACCCCAAAAAA 120593

QY 506 AAAGAA 511
Db 120592 GAAGAA 120587

RESULT 12
US-09-949-016-16554/c
; Sequence 16554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16554
; LENGTH: 168105
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168105)
; OTHER INFORMATION: n = A,T,C or G
```



```
/ LOCATION: (779676)..(779676)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (855539)..(855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (871619)..(871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1113881)..(1113881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349473)..(1349473)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349491)..(1349491)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1470091)..(1470091)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1569020)..(1569020)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1602912)..(1602912)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1603734)..(1603734)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1637998)..(1637998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1664854)..(1664855)
/ OTHER INFORMATION: n equals a, t, c, or g
/ US-08-916-421B-1
```

```
Query Match 6.2%; Score 44.4; DB 4; Length 1664976;
Best Local Similarity 52.6%; Pred. No. 1.1;
Matches 120; Conservative 0; Mismatches 106; Indels 2; Gaps 1;

QY 488 AAAAAAGAAAAGAAAAGAAAATAGCTTATATATAGCTATTATCTATCGTTATGT 547
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433470 AAAACAGAAAATGGTTATAGTTAGTAATATCTACATGGCTTTACATGGATGAAA 433529

QY 548 TTAGTTTTGGTAATAATAAGATCATCATATATGAATGTTGTGATCGTGT--AACTAA 605
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433530 TAAGTTTTTGCAACAAAACAAAATATTGTATGGATTGATTTTCGTATTTCAAAGAA 433589

QY 606 GGCAAGCTTAGGTATATAGACACCTTTAGAGTGCTTTTATGCGTGTGCTATGTTTGT 665
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433590 ACATCAATACAGATATTGAGGTATTATAAAGTAAGCAAGTCATGCTTTTGTAGTAT 433649

QY 666 TGCTGCAGAGTTGTAACCATCTTTGAAATATATAAAAAAGATCATGTTT 713
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 433650 AGTTGAACCATTTATACCCCTTCATATAGGAATAAAGTAAAGTATGTT 433697
```

```
RESULT 14
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
```


Qy 666 TGCTGCAGAGTTGTAAACCATCTTGAAATAATATAAAAAGATCATGTTT 713

Db 433650 AGTTGAACCATTTATACCACTTCATATAGGAATAAAGTAAAGATGTTT 433697

RESULT 15

```

US-09-949-016-14026/C
; Sequence 14026, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14026
; LENGTH: 58665
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14026

```

	Query Match	Best Local Similarity	6.2%;	Score 44.2;	DB 4;	Length 56665;
	Matches	8;	Conservative	0;	Mismatches	63;
					Indels	0;
					Gaps	0;
Qy	385	GGAGCTCAGAACTTGCCCTCAACAGTGC	CGCCCTTAGGGCAACACAGCGTTGGCATTCGA	444		
Db	31326	GGAAACAGGAGGTGAGGTTGCAGTGGACAGAGATCGCACCACCTGCACCTCCAGCTTCGG		31267		
Qy	445	CGTCTGAAGTGGCGGCAGACAGATCTAAACACCTATCTCAAAAAAGAAAGAAAG	504			
Db	31266	CGACAGAGTGTAGACTCCGTCTTCAGAAAAACAAACAAACAAACAAAAAAGAAAGAA	31207			

Search completed: August 24, 2005, 03:57:02
Job time : 152.962 sec8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 433.422 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 717
Sequence: 1 gctcacatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
 - 5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq*
 - 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq*
 - 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq*
 - 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq*
 - 22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq*
 - 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq*
 - 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq*
 - 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
 - 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	717	100.0	717	10	US-09-731-375A-2
2	717	100.0	717	20	US-10-728-323-2
3	717	100.0	717	20	US-10-728-051-2
4	717	100.0	717	21	US-10-899-551-3
5	474	66.1	474	14	US-10-228-806-3
6	474	66.1	474	17	US-10-100-303A-62
7	462	64.4	682	21	US-10-958-324-3

8	462	64.4	1162	21	US-10-958-324-1	Sequence 1, Appli
9	80	11.2	80	21	US-10-958-324-7	Sequence 7, Appli
10	65.6	9.1	1013	18	US-10-424-599-26383	Sequence 26383, A
11	63.4	8.8	770	17	US-10-302-633-1	Sequence 1, Appli
12	62	8.6	62	21	US-10-958-324-8	Sequence 8, Appli
13	54.4	7.6	883	18	US-10-424-599-121530	Sequence 121530,
14	49.8	6.9	62658	19	US-10-322-281-420	Sequence 420, App
15	49.6	6.9	52216	9	US-09-747-810-1	Sequence 1, Appli
16	49.6	6.9	169739	19	US-10-450-826-93	Sequence 93, Appl
17	49	6.8	349	18	US-10-424-599-133074	Sequence 133074,
18	49	6.8	62001	19	US-10-316-459-13	Sequence 13, Appl
19	48.2	6.7	7758	15	US-10-311-455-1076	Sequence 1076, Ap
20	48	6.7	6509	15	US-10-311-455-199	Sequence 199, App
21	47.8	6.7	1237	13	US-10-027-632-122633	Sequence 122633,
22	47.8	6.7	1237	17	US-10-027-632-122633	Sequence 122633,
23	47.6	6.6	451	18	US-10-424-599-24494	Sequence 24494, A
24	47.4	6.6	325	18	US-10-424-599-76496	Sequence 76496, A
25	46	6.4	1069	21	US-10-956-157-107	Sequence 107, App
26	46	6.4	1069	21	US-10-956-157-5342	Sequence 5342, Ap
27	45.6	6.4	6063	15	US-10-240-453-287	Sequence 287, App
28	45.4	6.3	5823	15	US-10-450-836-50	Sequence 256, App
29	45.2	6.3	121724	19	US-10-450-836-50	Sequence 50, Appl
30	45	6.3	48203	19	US-10-741-601-5711	Sequence 5711, Ap
31	45	6.3	48203	21	US-10-741-600-17815	Sequence 17815, A
32	44.4	6.2	6385	14	US-10-239-676-58	Sequence 58, Appl
33	44.4	6.2	6385	15	US-10-240-453-60	Sequence 60, Appl
34	44	6.1	812	13	US-10-027-632-23973	Sequence 23973, A
35	44	6.1	812	17	US-10-027-632-23973	Sequence 23973, A
36	43.8	6.1	7566	10	US-09-764-891-10010	Sequence 10010, A
37	43.8	6.1	174448	13	US-10-087-192-148	Sequence 148, App
38	43.6	6.1	60815	13	US-10-087-192-52	Sequence 52, Appl
39	43.4	6.1	5919	15	US-10-311-455-362	Sequence 362, App
40	43.4	6.1	5919	17	US-10-221-613-64	Sequence 64, Appl
41	43.4	6.1	19696	10	US-09-764-891-9327	Sequence 9327, Ap
42	43.4	6.1	19696	14	US-10-091-572-874	Sequence 874, App
43	43.4	6.1	24173	10	US-09-764-891-9294	Sequence 9294, Ap
44	43.4	6.1	24173	14	US-10-091-572-855	Sequence 855, App
45	43.2	6.0	151	18	US-10-424-599-107242	Sequence 107242,

ALIGNMENTS

RESULT 1
US-09-731-375A-2
; Sequence 2, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731,375A
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-731-375A-2

Query Match	100.0%	Score 717;	DB 10;	Length 717;
Best Local Similarity	100.0%	Pred. No. 2.4e-178;		
Matches 717;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCTCACATACTAGTAGCCCTCGCCCTTTCTCTCGTGCACGATCTGCGAGGCA	60	
Db	1	GCTCACATACTAGTAGCCCTCGCCCTTTCTCTCGTGCACGATCTGCGAGGCA	60	
Qy	61	GCAGTGGGAATCCACAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGACCTGAG	120	

Db 61 GCAGTGGGAACCTCCAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTGTAG 120
QY 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAACTGTGAGAGGATTCATATCAACGGGA 180
Db 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAACTGTGAGAGGATTCATATCAACGGGA 180
QY 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
QY 241 ATCTCTCAGCACACAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 241 ATCTCTCAGCACACAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
QY 301 GTGCATGTGCGAGGCAATTCACACAGATCATGGAGAACCAAGAGAGATGTTGCGAGGGAG 360
Db 301 GTGCATGTGCGAGGCAATTCACACAGATCATGGAGAACCAAGAGAGATGTTGCGAGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
QY 421 GGCACACAGCGTTGCGACTTGCAGCTCGAAAGTGGCGGAGAGACAGATACCTAAACACC 480
Db 421 GGCACACAGCGTTGCGACTTGCAGCTCGAAAGTGGCGGAGAGACAGATACCTAAACACC 480
QY 481 TATCTCAAAAAAGAAAAAGAAAGAAATAGCTTATATAGCTTATATCTATG 540
Db 481 TATCTCAAAAAAGAAAAAGAAAGAAATAGCTTATATAGCTTATATCTATG 540
QY 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCACTATATGATGTTGTCGTGTTA 600
Db 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCACTATATGATGTTGTCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCTTTATGCGGCTGTCTATGT 660
Db 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCTTTATGCGGCTGTCTATGT 660
QY 661 TTTGTTGCTCGAGAGTTGTAACCATCTTGAATAATATAAAAAAGATCATGTTTGT 717
Db 661 TTTGTTGCTCGAGAGTTGTAACCATCTTGAATAATATAAAAAAGATCATGTTTGT 717

RESULT 2

US-10-728-323-2
; Sequence 2, Application US/10728323
; Publication No. US20040208894A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0232
; CURRENT APPLICATION NUMBER: US/10/728,323
; PRIOR FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea

US-10-728-323-2

Query Match 100.0%; Score 717; DB 20; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.4e-178;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCTCACCATATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCACTGCGAGGCA 60
Db 1 GCTCACCATATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCACTGCGAGGCA 60
QY 61 GCAGTGGGAACCTCCAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTGTAG 120
Db 61 GCAGTGGGAACCTCCAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACTGTAG 120
QY 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAACTGTGAGAGGATTCATATCAACGGGA 180
Db 121 GCCTCGGAGCAACATCTCATGACAGAGATCCAACTGTGAGAGGATTCATATCAACGGGA 180
QY 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
Db 181 CCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGG 240
QY 241 ATCTCTCAGCACACAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
Db 241 ATCTCTCAGCACACAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACAAAG 300
QY 301 GTGCATGTGCGAGGCAATTCGAAACAGATCATGAGAACCAAGAGAGATGTTGCGAGGGAG 360
Db 301 GTGCATGTGCGAGGCAATTCGAAACAGATCATGAGAACCAAGAGAGATGTTGCGAGGGAG 360
QY 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
Db 361 GCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAGTGGCGCTTAG 420
QY 421 GGCACACAGCGTTGCGACTTGCAGCTCGAAAGTGGCGGAGAGACAGATACCTAAACACC 480
Db 421 GGCACACAGCGTTGCGACTTGCAGCTCGAAAGTGGCGGAGAGACAGATACCTAAACACC 480
QY 481 TATCTCAAAAAAGAAAAAGAAAGAAATAGCTTATATAGCTTATATCTATG 540
Db 481 TATCTCAAAAAAGAAAAAGAAAGAAATAGCTTATATAGCTTATATCTATG 540
QY 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCACTATATGATGTTGTCGTGTTA 600
Db 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCACTATATGATGTTGTCGTGTTA 600
QY 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCTTTATGCGGCTGTCTATGT 660
Db 601 ACTAAGCAAGCTTAGTTATATGAGCACTTTAGAGTGCTTTATGCGGCTGTCTATGT 660
QY 661 TTTGTTGCTCGAGAGTTGTAACCATCTTGAATAATATAAAAAAGATCATGTTTGT 717
Db 661 TTTGTTGCTCGAGAGTTGTAACCATCTTGAATAATATAAAAAAGATCATGTTTGT 717

RESULT 3

US-10-728-051-2
; Sequence 2, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10/728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Arachis hypogaea

US-10-728-051-2

Query Match 100.0%; Score 717; DB 20; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.4e-178;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACCGCATCTCGAGGCA 60
Db 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACCGCATCTCGAGGCA 60

Qy 61 GCAGTGGGAATCTCAAGGAGACAGAAGATGTCAGAGCCAGTCCGAGAGGCGCAACCTGAG 120
Db 61 GCAGTGGGAATCTCAAGGAGACAGAAGATGTCAGAGCCAGTCCGAGAGGCGCAACCTGAG 120

Qy 121 GCCCTGCGAGCAATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180
Db 121 GCCCTGCGAGCAATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180

Qy 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 240
Db 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 240

Qy 241 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
Db 241 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300

Qy 301 GTGCATGTGGAGGATTCGCAACAGATCATGCAAGATCATGCAAGATCATGCAAGATCATGCAAG 360
Db 301 GTGCATGTGGAGGATTCGCAACAGATCATGCAAGATCATGCAAGATCATGCAAGATCATGCAAG 360

Qy 361 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 420
Db 361 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 420

Qy 421 GGCACCAACAGCGTTGCGACTTGGAAGTGGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 480
Db 421 GGCACCAACAGCGTTGCGACTTGGAAGTGGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 480

Qy 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540

Qy 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCATATGAAATGTTGATCGTGTTA 600
Db 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCATATGAAATGTTGATCGTGTTA 600

Qy 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Db 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGCTATGT 660

Qy 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 717

RESULT 4

US-10-899-551-3
; Sequence 3, Application US/10899551
; Publication No. US2005006394A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 717

; TYPE: DNA
; ORGANISM: species Arachis hypogaea
; US-10-899-551-3

Query Match 100.0%; Score 717; DB 21; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.4e-178;
Matches 717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACCGCATCTCGAGGCA 60
Db 1 GCTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACCGCATCTCGAGGCA 60

Qy 61 GCAGTGGGAATCTCAAGGAGACAGAAGATGTCAGAGCCAGTCCGAGAGGCGCAACCTGAG 120
Db 61 GCAGTGGGAATCTCAAGGAGACAGAAGATGTCAGAGCCAGTCCGAGAGGCGCAACCTGAG 120

Qy 121 GCCCTGCGAGCAATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180
Db 121 GCCCTGCGAGCAATCTCATGCAAGAGATCCAACTGACGAGGATTCATATGAACGGGA 180

Qy 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 240
Db 181 CCCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGG 240

Qy 241 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300
Db 241 ATCCTCTCAGCACCAAGAGAGGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAG 300

Qy 301 GTGCATGTGGAGGATTCGCAACAGATCATGCAAGATCATGCAAGATCATGCAAGATCATGCAAG 360
Db 301 GTGCATGTGGAGGATTCGCAACAGATCATGCAAGATCATGCAAGATCATGCAAGATCATGCAAG 360

Qy 361 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 420
Db 361 GCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 420

Qy 421 GGCACCAACAGCGTTGCGACTTGGAAGTGGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 480
Db 421 GGCACCAACAGCGTTGCGACTTGGAAGTGGAGGAGCTCAGGAACTTGCCCTCAACAGTGCGGCCCTTAG 480

Qy 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540
Db 481 TATCTCAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 540

Qy 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCATATGAAATGTTGATCGTGTTA 600
Db 541 GTTATGTTTGTAGTTTGGTAATAAAGATCATCATATGAAATGTTGATCGTGTTA 600

Qy 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGCTATGT 660
Db 601 ACTAAGGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTTTTATGGCGTTGCTATGT 660

Qy 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 717
Db 661 TTTGTTGCTGCAGAGTTGTAACCATCTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 717

RESULT 5

US-10-228-806-3
; Sequence 3, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 474
; TYPE: DNA

; ORGANISM: Arachis hypogaea		; ORGANISM: Arachis hypogaea	
US-10-228-806-3		US-10-958-324-3	
Query Match 66.1%; Score 474; DB 14; Length 474;		Query Match 64.4%; Score 462; DB 21; Length 682;	
Best Local Similarity 100.0%; Pred. No. 2e-114;		Best Local Similarity 100.0%; Pred. No. 3.6e-111;	
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 61	QY 1	GCTCACCATACATAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60
DB 1	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60	DB 67	GCTCACCATACATAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 126
QY 62	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 121	QY 61	GCAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAG 120
DB 61	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 120	DB 127	GCAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAG 186
QY 122	CCCTGGGAGCAACATCTCATGCGAGAGATCCACGTCGACGAGATTCATATGAACGGGAC 181	QY 121	GCCTCGCGAGCAACATCTCATGCGAGAGATCCAAAGATTCATATGAACGGGAC 180
DB 121	CCCTGGGAGCAACATCTCATGCGAGAGATCCACGTCGACGAGATTCATATGAACGGGAC 180	DB 187	GCCTCGCGAGCAACATCTCATGCGAGAGATCCAAAGATTCATATGAACGGGAC 246
QY 182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 241	QY 181	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 240
DB 181	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 240	DB 247	CCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 306
QY 302	TGATGTGGAGGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAG 361	QY 241	ATCTCTCTCAGCAGCAAGAGAGAGGTTGCAATCAGCTGCAACGAGTTTGAGAAACCAACCAAG 300
DB 301	TGATGTGGAGGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAGATTCGAAAG 360	DB 307	ATCTCTCTCAGCAGCAAGAGAGAGGTTGCAATCAGCTGCAACGAGTTTGAGAAACCAACCAAG 366
QY 362	CAACAGGAGCAACAGTTTCAAGAGGAGCTCAGAACTTCGCTCAACAGTGCAGGCTTAGG 421	RESULT 7	
DB 361	CAACAGGAGCAACAGTTTCAAGAGGAGCTCAGAACTTCGCTCAACAGTGCAGGCTTAGG 420	US-10-958-324-3	
QY 422	GCACCACAGCGTTGCGACTTGGACCTCGAAAGTGGCGGCGAGACAGATACATAA 475	; Sequence 3, Application US/10958324	
DB 421	GCACCACAGCGTTGCGACTTGGACCTCGAAAGTGGCGGCGAGACAGATACATAA 474	; Publication No. US20050114924A1	
GENERAL INFORMATION:		GENERAL INFORMATION:	
; APPLICANT: DODO, HORTENSE W.		; APPLICANT: DODO, HORTENSE W.	
; APPLICANT: ARNTZEN, CHARLES J.		; APPLICANT: ARNTZEN, CHARLES J.	
; APPLICANT: KONAN, KOFFI N'DA		; APPLICANT: KONAN, KOFFI N'DA	
; APPLICANT: VIQUEZ, OLGA		; APPLICANT: VIQUEZ, OLGA	
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN		; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN	
; FILE REFERENCE: 072121/0104		; FILE REFERENCE: 072121/0104	
; CURRENT APPLICATION NUMBER: US/10/958,324		; CURRENT APPLICATION NUMBER: US/10/958,324	
; CURRENT FILING DATE: 2004-10-06		; CURRENT FILING DATE: 2004-10-06	
; PRIOR APPLICATION NUMBER: US/09/715,036		; PRIOR APPLICATION NUMBER: US/09/715,036	
; PRIOR FILING DATE: 2000-11-20		; PRIOR FILING DATE: 2000-11-20	
; PRIOR APPLICATION NUMBER: 60/167,255		; PRIOR APPLICATION NUMBER: 60/167,255	
; PRIOR FILING DATE: 1999-11-19		; PRIOR FILING DATE: 1999-11-19	
; NUMBER OF SEQ ID NOS: 8		; NUMBER OF SEQ ID NOS: 8	
; SOFTWARE: Patent In Ver. 2.1		; SOFTWARE: Patent In Ver. 2.1	
; SEQ ID NO 3		; SEQ ID NO 3	
; LENGTH: 682		; LENGTH: 682	
; TYPE: DNA		; TYPE: DNA	
; ORGANISM: Arachis hypogaea		; ORGANISM: Arachis hypogaea	
US-10-958-324-3		US-10-100-303A-62	
Query Match 66.1%; Score 474; DB 17; Length 474;		Query Match 66.1%; Score 474; DB 17; Length 474;	
Best Local Similarity 100.0%; Pred. No. 2e-114;		Best Local Similarity 100.0%; Pred. No. 2e-114;	
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 61	QY 2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 61
DB 1	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60	DB 1	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCACTCGGAGGCAG 60
QY 62	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 121	QY 62	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 121
DB 61	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 120	DB 61	CAGTGGGAATCCAAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGAACTTGAGG 120
QY 122	CCCTGGGAGCAACATCTCATGCGAGAGATCCACGTCGACGAGATTCATATGAACGGGAC 181	QY 122	CCCTGGGAGCAACATCTCATGCGAGAGATCCACGTCGACGAGATTCATATGAACGGGAC 181

Qy 301 GTGCATGTGCGAGGCAATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Db 367 GTGCATGTGCGAGGCAATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 426
Qy 361 GCAACAGGAGCAACAGATTCAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCCTTAG 420
Db 427 GCAACAGGAGCAACAGATTCAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCCTTAG 486
Qy 421 GGCACACAGCGTTGCGACTTGGACGTCGCGAAAGTGGCGGCAG 462
Db 487 GGCACACAGCGTTGCGACTTGGACGTCGCGAAAGTGGCGGCAG 528

RESULT 8
US-10-958-324-1
; Sequence 1, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (110)..(730)
US-10-958-324-1

Query Match 64.4%; Score 462; DB 21; Length 1162;
Best Local Similarity 100.0%; Pred. No. 4.7e-111;
Matches 462; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCAATCTGCGAGGCA 60
Db 118 GCTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCAATCTGCGAGGCA 177
Qy 61 GCAGTGGAACTCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGGAACCTTAG 120
Db 178 GCAGTGGAACTCCAAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGGAACCTTAG 237
Qy 121 GCCTGCGAGCAATCTCATGAGAGATCCACGTCGAGGAGGATTCATATGACGGGA 180
Db 238 GCCTGCGAGCAATCTCATGAGAGATCCACGTCGAGGAGGATTCATATGACGGGA 297
Qy 181 CCGGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCCATATATGATCGGAGGCGCTGG 240
Db 298 CCGGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCCATATATGATCGGAGGCGCTGG 357
Qy 241 ATCTCTCAGACCAAGAGAGGTTTGCATATGAGTGAACGAGTTTGAGAACCAACCAAG 300
Db 358 ATCTCTCAGACCAAGAGAGGTTTGCATATGAGTGAACGAGTTTGAGAACCAACCAAG 417
Qy 301 GTGCATGTGCGAGGCAATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 360
Db 418 GTGCATGTGCGAGGCAATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCGAGGGAG 477
Qy 361 GCAACAGGAGCAACAGATTCAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCCTTAG 420
Db 478 GCAACAGGAGCAACAGATTCAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCCTTAG 537

Qy 421 GGCACACAGCGTTGCGACTTGGACGTCGCGAAAGTGGCGGCAG 462
Db 538 GGCACACAGCGTTGCGACTTGGACGTCGCGAAAGTGGCGGCAG 579

RESULT 9
US-10-958-324-7
; Sequence 7, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-958-324-7

Query Match 11.2%; Score 80; DB 21; Length 80;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCAATCTCGAGGCGAGCGAGTGGAA 70
Db 1 CTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCAATCTCGAGGCGAGCGAGTGGAA 60

Qy 71 CTCACAGGAGACAGAGATG 90
Db 61 CTCACAGGAGACAGAGATG 80

RESULT 10
US-10-424-599-26383
; Sequence 26383, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 26383
; LENGTH: 1013
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1013)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123825C.1
US-10-424-599-26383

Query Match		9.1%;	Score 65.6;	DB 18;	Length 1013;	
Best Local Similarity		50.8%;	Pred. No. 1.2e-06;			
Matches 363;		Conservative 0;	Mismatches 309;	Indels 42;	Gaps 7;	
QY	7	CATACTAGTAGCCCTCGCCCTTTCTCTCTCGTGCCTCCAGCATCTGCGAGGAGCAGTG	66			
DB	76	CACAATCCTCCTCATCTCTCTCTCTCTGTCATCGCCACACTTGACGCGCTCCAAATG	135			
QY	67	GGAACTCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGCAACTGAGGCCCTG	126			
DB	136	GCAGCACCAACAAGA---TAGCTGCCGAAGACAGCTCCAGGGGGTGAACCTCAGCCCTG	192			
QY	127	CGAGCAACATCTCATGCAGAAAGATCCAACTGACGAGGATTCTATTAAGCGGACCCGTA	186			
DB	193	CGAGAGCACATCATGAGAGATCCAGGCCCGCGCATGACGATGATGATGATGACGA	252			
QY	187	CAGC-----CCTAGTCAGGATCCGTACAGCCCTAGTTCATATGATCGGAG-----	231			
DB	253	CGACAATCATCTCAGGACCATCGGGGAAGAATCACTATAAGGAGGAACGAAGG	312			
QY	232	-----AGGCGCTGCCTCTCAGCACCAAGAGGTGTGCAATGAGCTGAACGA	282			
DB	313	AAAAGACGAAGACGAAGAAGAGACACATGCAGAAGTGTGCACAGAAATGAGCGA	372			
QY	283	GTTTGAGAACAAACAAAGGTGCATGTCCGAGGCAATTGCCACAGATCATGGAGAACGAG	342			
DB	373	GCT---GAGAGGCCCAATGCCAGTGCAGAGCGCTGCAGAGATTAATGGAGAACGAG	429			
QY	343	CGATAGTTCAGGGAGGCAACAGAGCAACAGTTCAGAGGGAGCTCAGGAATTTGCC	402			
DB	430	CGAGGAATCTGGAGAGAGCAGAAGAAGAAA---ATGGAGAGGAGCTCATTAATTTGGC	486			
QY	403	TCAACAGTGGCGCTTAGGGCACCAACAGCGTTGGAGCTTGCAAGTGGCGGCGAG	462			
DB	487	TACTATGTGAGGTTTGGAGCCCATGATCCAGTGGGACTTGTCTCCGA-----TGACTA	540			
QY	463	AGACAGATACTAAACACCTATCTCAAAAAAAGAAAGAAAGAAAGAAATAGCTTATA	522			
DB	541	AGAAITTAAGCAATGTCTACTTGTAGTACTAACACATGATGTGATGTTATGCT	600			
QY	523	TATAAGTATATCTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATG	582			
DB	601	AGTAGTATAAATAAGCTGTCTCTGAGTGTGTGTATATTAATAAAGATCATCACTGG	660			
QY	583	AATGTGTGATCTGTAACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCTT	642			
DB	661	TGAATGGTGTGTTGACGTACCTACTTAGTAGGCAATGGAAGCACTTAGAGTGTGCTT	720			
QY	643	T---TATGGCTGTCTATGTTTGTGTCGAGAGTTGTAACCATCTTGAAT	693			
DB	721	TGTGCAATGGCTTGCCTCTGTTTGAGACTTTTGTAAATGTTTCGAGTTAAAT	774			
RESULT 11						
US-10-302-633-1						
; Sequence 1, Application US/10302633						
; Publication No. US20030229038A1						
; GENERAL INFORMATION:						
; APPLICANT: de Lumen, Benito O.						
; Galvez, Alfredo F.						
; TITLE OF INVENTION: Lunasin Peptides						
; NUMBER OF SEQUENCES: 3						
; CORRESPONDENCE ADDRESS:						
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP						
; STREET: 75 DENISE DRIVE						
; CITY: HILLSBOROUGH						
; STATE: CALIFORNIA						
; COUNTRY: USA						
; ZIP: 94010						
; COMPUTER READABLE FORM:						
; MEDIUM TYPE: Floppy disk						
; COMPUTER: IBM PC compatible						

OPERATING SYSTEM: PC-DOS/MS-DOS						
SOFTWARE: PatentIn Release #1.0, Version #1.30						
CURRENT APPLICATION DATA:						
APPLICATION NUMBER: US/10/302,633						
FILING DATE: 22-No. US20030229038A1-2002						
CLASSIFICATION: <Unknown>						
PRIOR APPLICATION DATA:						
APPLICATION NUMBER: US/09/531,727						
FILING DATE: 21-Mar-2000						
APPLICATION NUMBER: 08/939,675						
FILING DATE: <Unknown>						
ATTORNEY/AGENT INFORMATION:						
NAME: OSMAN, RICHARD A						
REGISTRATION NUMBER: 36,627						
REFERENCE/DOCKET NUMBER: B98-003						
TELECOMMUNICATION INFORMATION:						
TELEPHONE: (650) 343-4341						
TELEFAX: (650) 343-4342						
INFORMATION FOR SEQ ID NO: 1:						
SEQUENCE CHARACTERISTICS:						
LENGTH: 770 base pairs						
TYPE: nucleic acid						
STRANDEDNESS: double						
TOPOLOGY: linear						
MOLECULE TYPE: cdna						
SEQUENCE DESCRIPTION: SEQ ID NO: 1:						
US-10-302-633-1						
Query Match		8.8%;	Score 63.4;	DB 17;	Length 770;	
Best Local Similarity		51.6%;	Pred. No. 4e-06;			
Matches 369;		Conservative 0;	Mismatches 301;	Indels 45;	Gaps 8;	
QY	7	CATACTAGTAGCCCTCGCCCTTTTCTCTCTCGTGCCTCCACGCACTCTGCGAGGAGCAGTG	66			
DB	28	CACAATCTCTCTCACTCTCTCTCTCTCTGTCATGCCACACTTGCAGCGCTCCAAATG	87			
QY	67	GGAACTCCAAAGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGGCAACTGAGGCCCTG	126			
DB	88	GCAGCACCAAGCAAGA---TAGCTGCCGAAGCAGCTCCAGGGGGTGAACCTCAGCGCCTG	144			
QY	127	CGAGCAACATCTCATGACAGAAGATCCAACTGACGAGGATTCATATGAACGGGACCCGTA	186			
DB	145	CGAGAAGCACTCATGAGAGAAGATCCAGGCCCGCGCGATGACGATGATGATGATGACGA	204			
QY	187	CAGC-----CCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAG-----	231			
DB	205	CGACAATCACTCTCAGGACCATCGGGGGAAGATCACTACATGAAGAGGAGCAAGG	264			
QY	232	-----AGGCGCTGGATCCTCTCAGCAACCAAGAGAGGTGTTGCAATGAGCTGAACGA	282			
DB	265	AAAAGACGAAGACGAAGAAGAGAGGACACATGCAGAAAGTGTGTCACAGAAATGAGCGA	324			
QY	283	GTTTGAGAACAAACAAAGGTGCATGTGCGAGGCAATGCAACAGATCATCGAGAACGAG	342			
DB	325	GCT---GAGAAGCCCCAAATGCCAGTGCAGAAAGCGCTGCAGAAAGATTAATGGAGAACGAG	381			
QY	343	CGATAGTTGTCAGGGGAGGCAACAGAGCAACAGATTTCAGAGGGAGCTCAGGAACCTTGCC	402			
DB	382	CGAGGAACTGGAGGAGACAGACAGAGAA---ATGGAGAGGAGCTCATTAACCTTGGC	438			
QY	403	TCAACAGTGGCGCTTAGGGCACCAACAGCGTTGCGACTTGGAGCGTGA---AAGTGGCGGCA	461			
DB	439	TACTATGTGAGGTTTGGACCCCATGATCCAGTGCAGCTTGTCTCTCGATGACTAAGAAGT	498			
QY	462	GAGACAGATACTAAACACCTATCTCAAAAAAGAAAGAAAGAAAGAAATAGCTTAT	521			
DB	499	TAAAAGCAATGTGTCACTTGTCTGCTACTAACACATGATGTGATGTTTATGCTAGCT---	555			
QY	522	ATATAAGCTATTATCTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTAT	581			
DB	556	-----AGCTATAACATAAGCTGTCTGAGTGTGTTGTATATTAATAAAGATCATCACTG	610			
QY	582	GAATGTGTGATCGTGTAACTAAGCAAGCTTAGGTTATATGAGCACCTTTAGAGTGCT	641			

Db 611 GTGAATGGTATCGTGTACGTACCTACTAGTAGGCAATGGAAGCACTTAGAGTGTGCT 670
Qy 642 TT---TATGGCGTGTCTATGTTTGTGTGTCGACAGAGTTGTAACCATCTTTGAAAT 693
Db 671 TTGTGCATGGCGCTTGCCTCTGTTTGTGAGACTTTTGTAAATGTTTTCGAGTTTAAAT 725

RESULT 12

US-10-958-324-8
; Sequence 8, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIQUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 62
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Probe
US-10-958-324-8

Query Match 8.6%; Score 62; DB 21; Length 62;

Best Local Similarity 100.0%; Pred. No. 2.7e-06; Indels 0; Gaps 0;
Matches 62; Conservative 0; Mismatches 0;

Qy 301 GTGCATGTGGCGAGCATTCGCAACAGATCATGGAGAACGACGATAGTTGACGGGGAG 360
Db 1 GTGCATGTGGCGAGCATTCGCAACAGATCATGGAGAACGACGATAGTTGACGGGGAG 60
Qy 361 GC 362
Db 61 GC 62

RESULT 13

US-10-424-599-121530
; Sequence 121530, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121530
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80749C.1
US-10-424-599-121530

Query Match 7.6%; Score 54.4; DB 18; Length 883;

Best Local Similarity 61.5%; Pred. No. 0.001;
Matches 123; Conservative 0; Mismatches 71; Indels 6; Gaps 2;
Qy 251 CACCAAGAGAGAGTGTGTCATGAGCTGAACGAGTTTGAGAACCAACAAAGGTCATGTC 310
Db 437 CACATGCAGAAGTGTCTGCAGCGAAATGAGCGAGCT---GAAAGCCCATATGCCAGTGC 493
Qy 311 GAGGCATTGCCAACACAGATCATGGAGAACGACGAGGATAGTTGAGGGGAGGCAACAGAG 370
Db 494 AAAGCGCTACAGAAGATAATGGAATAACGAGGCGAGCAACTGGAGGGGAAGGAGAAG 553
Qy 371 CAACAGTTCAAGAGGAGGAGCTCAGGAACCTGCCCTCAACAGTGGCGGCTTTAGGGACACAG 430
Db 554 ---CAGATGGAGAGAGAGCTCATGAACCTTGGCTATTAGTGCAGGTTGGGACCATGATA 610
Qy 431 CGTTGGCGACTTGGACGTCGA 450
Db 611 GGGTGGCGACTTGTCTCCGA 630

RESULT 14

US-10-322-281-420/c
; Sequence 420, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 420
; LENGTH: 62658
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-420

Query Match 6.9%; Score 49.8; DB 19; Length 62658;

Best Local Similarity 54.1%; Pred. No. 0.14;
Matches 125; Conservative 0; Mismatches 102; Indels 4; Gaps 1;

Qy 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGCGTTCCGACTTGGAC 445
Db 61081 GAGCCCGAGGAGTTGAGGGCTGCAGTGAGCCATGATTCACCACTGTCTCCAGCTGGGC 61022
Qy 446 GTCGAAAGTGGCGGACAGACAGATATAACACCTATCTCAAAAAAGAAAGAAAGA 505
Db 61021 GACAGAGCAAGACTCTGTCTCAAAAAAGAAAGAAAGAAAGAAAGA 60962
Qy 506 AAAGAAAATAGCTTATATATAAGCTATTATCTATGTTATGTTTGTGTAATAATA 565
Db 60961 AAAGAAAAGGA---AAATAAGCTAATGCTTGTGTTATGGGAGTCAAGGACACACA 60906
Qy 566 AAGATCATCATATATGAATGTGTTGATCGTGTAACTAAGGCAAGCTTAG 616
Db 60905 GTGGCAAGGTAAGTTCAATCTTCAGTTAGTTTGGCAAGTTTAAGTTTG 60855

RESULT 15

US-09-747-810-1/c
; Sequence 1, Application US/09747810
; Patent No. US20020012903A1
; GENERAL INFORMATION:
; APPLICANT: Okano, Hideyuki
; APPLICANT: Goldman, Steven A.
; TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGENI
; FILE REFERENCE: 19603/3580
; CURRENT APPLICATION NUMBER: US/09/747,810
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,003

```
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 52216
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-747-810-1

Query Match      6.9%; Score 49.6; DB 9; Length 52216;
Best Local Similarity 61.7%; Pred. No. 0.15; 49; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 386 GAGCTCAGGAACCTTGCCCTCAACAGTGGCGCTTTAGGGCCACACAGCGTTGCGACTTGGAC 445
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12026 GAACTCAGGAGCTGGAGTTTGCAAGTGGCGAGATCGCACCACCTGCATTCCAGCCTGGGC 11967

QY 446 GTCGAAAGTGGCGGCGAGACAGATACTAAACACCTATCTCAAAAAAGAAAAAGAAAGA 505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11966 AACAGAGTGAGACTCTGTCTCAAAAAATAAAAAAAATTTAAAAAGAAAAAGAAAGA 11907

QY 506 AAAGAAAA 513
    ||||| |||||
Db 11906 AAAAGAA 11899
```

Search completed: August 24, 2005, 09:47:13
Job time : 435.422 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 40.6071 Seconds
(without alignments)
13658.075 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DRV=xlh
-Q=/cgn2.1/USPTO spo01/US10728323/runat 23082005 124355 29200/app_query.fasta 1.4757
-DB=A Geneseq 16Dec04 -OPMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1 224 @runat 23082005 124355 29200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 16Dec04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	66.6	157	2 AAW24153	Aaw24153 Peanut al
2	842	66.6	157	2 AAW24164	Aaw24164 Peanut al
3	842	66.6	157	2 AAY15245	Aay15245 Peanut al
4	842	66.6	157	4 AAU04707	Aau04707 Anaphylac
5	842	66.6	157	6 ABUS2464	Abu52464 Peanut Ar
6	842	66.6	157	7 ADG27518	Adg27518 Peanut al
7	838.5	66.3	207	4 AAB82383	Aab82383 Peanut al
8	830	65.6	156	2 AAY40973	Aay40973 Ara h 2 p
9	819	64.7	156	8 ADO38314	Ado38314 Ara h 2 M
10	814	64.3	157	6 ABUS2576	Abu52576 Peanut Ar

11	812	64.2	156	8 ADM12096	Adm12096 Arachis h
12	812	64.2	157	6 ABUS2575	Abu52575 Peanut Ar
13	774	61.2	157	6 ABUS2577	Abu52577 Peanut Ar
14	771	60.9	166	2 AAY40968	Aay40968 Recombina
15	771	60.9	166	6 ABUS2482	Abu52482 Peanut Ar
16	771	60.9	166	7 ADG27536	Adg27536 T7/His-ta
17	699	55.3	166	3 AAB33600	Aab33600 Modified
18	699	55.3	166	4 AAU05035	Aau05035 Modified
19	688.5	54.4	167	4 AAU04710	Aau04710 Modified
20	243	19.2	158	2 AAW23419	Aaw23419 Soybean a
21	243	19.2	158	7 AAY05723	Aay05723 Soybean G
22	243	19.2	158	7 ADH89269	Adh89269 G. max 2S
23	243	19.2	158	8 ADG44004	Adg44004 G. max 2S
24	241.5	19.1	155	2 AAW23418	Aaw23418 Soybean a
25	241.5	19.1	155	7 ADH89267	Adh89267 G. max na
26	241.5	19.1	155	8 ADG44002	Adg44002 G. max 2S
27	229	18.1	158	2 AAW23420	Aaw23420 Chimeric
28	184.5	14.6	323	7 ADH89277	Adh89277 Sunflower
29	184.5	14.6	323	8 ADG44012	Adg44012 H. annuus
30	177.5	14.0	168	6 AAE36076	Aae36076 Flax Conl
31	156	12.3	28	2 AAW24194	Aaw24194 Peanut al
32	156	12.3	158	2 AAW23586	Aaw23586 Mabinlin
33	151	11.9	158	2 AAW23588	Aaw23588 Mabinlin
34	151	11.9	295	7 ADH89275	Adh89275 Sunflower
35	151	11.9	295	8 ADG44010	Adg44010 H. annuus
36	149.5	11.8	167	8 ADS16393	Ads16393 Grape 2S
37	146.5	11.6	164	1 AAP91892	Aap91892 1kb fragm
38	146.5	11.6	164	1 AAP96144	Aap96144 Sequence
39	146.5	11.6	164	7 ADH89227	Adh89227 A. thalia
40	146.5	11.6	164	8 ADG43962	Adg43962 A. thalia
41	144	11.4	169	6 AAE36077	Aae36077 Flax Conl
42	140.5	11.1	146	6 AAO27313	Aao27313 Brazil nu
43	140.5	11.1	146	8 ADN37294	Adn37294 Brazil nu
44	139	11.0	140	1 AAP91891	Aap91891 Brazil nu
45	139	11.0	155	2 AAW23587	Aaw23587 Mabinlin

ALIGNMENTS

RESULT 1

AAW24153

ID AAW24153 standard; protein; 157 AA.

AC AAW24153;

DT 17-OCT-2003 (revised)

DT 29-DEC-1997 (first entry)

DE Peanut allergen Ara hII.

XX Peanut; seed storage protein; allergen; hypersensitivity;

KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;

KW ELISA; analysis; Ara hII.

XX Arachis hypogaea; strain Florunner.

XX W09724139-A1.

PD 10-JUL-1997.

XX 23-SEP-1996; 96WO-US015222.

XX 29-DEC-1995; 95US-0009455P.

PR 04-MAR-1996; 96US-00610424.

(UYAR-) UNIV ARKANSAS.

PI Burks AW, Helm RM, Cockrell G, Stanley JS, Bannan GA;

XX WPI; 1997-363453/33.

DR N-PSDB; AAT76615.

XX Peanut allergens Ara hI and Ara hII - used for vaccination and in two-

PT

PT	site monoclonal antibody based ELISA.	KW	vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
XX		KW	ELISA; analysis; Ara hII.
PS		XX	
XX	Claim 31; Page 198; 354pp; English.	OS	Arachis hypogaea; strain Florunner.
CC	This polypeptide comprises major peanut allergen Ara hII. Its sequence	XX	
CC	was deduced from a cDNA clone (AA76615) isolated from peanut seed cDNA	PN	WO9724139-A1.
CC	using a primer (see AAT76617) based on an isolated Ara hI peptide (see	PD	
CC	AAW24151). The sequence shows significant homology with the conglutin	XX	10-JUL-1997.
CC	family of seed storage proteins of other legumes. The allergen is	XX	
CC	recognised by serum IgE from a large proportion of individuals with	PF	23-SEP-1996; 96WO-US015222.
CC	peanut hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used	XX	
CC	to raise monoclonal antibodies which are used in a specific two-site MAb	XX	
CC	ELISA for the detection of Ara hI or Ara hII (claimed). IgE- binding Ara	PR	29-DEC-1995; 95US-0009455P.
CC	hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect	PR	04-MAR-1996; 96US-00610424.
CC	against allergic reactions to peanut allergens, e.g. anaphylactic shock.	XX	
CC	(Updated on 17-OCT-2003 to standardise OS field)	XX	(UYAR-) UNIV ARKANSAS.
XX		XX	
XX	Sequence 157 AA;	PI	Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
SQ		XX	
	Alignment Scores:	DR	WPI; 1997-363453/33.
	Pred. No.: 8.2e-90 Length: 157	DR	N-PSDB; AAT76614.
	Score: 842.00 Matches: 157	XX	
	Percent Similarity: 100.00% Conservative: 0	XX	Peanut allergens Ara hI and Ara hII - used for vaccination and in two-
	Best Local Similarity: 100.00% Mismatches: 0	PT	site monoclonal antibody based ELISA.
	Query Match: 66.56% Indels: 0	XX	
	DB: 2 Gaps: 0	PS	Claim 31; Page 219; 354pp; English.
		XX	
US-10-728-323-2 (1-717) x AAW24153 (1-157)		XX	This polypeptide comprises major peanut allergen Ara hII. Its sequence
QY	2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGTCCACGATCTGCGAGGAG 61	CC	was deduced from cDNA clone P38 (AAT76614), isolated from peanut seed
Db	1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20	CC	cDNA using a primer (see AAT76617) based on an isolated Ara hI peptide
QY	62 CAGTGGGAATCCCAAGGAGACAGAGATGCCAGGCAGCTCCAGAGGGCGACCTGAGG 121	CC	(see AAW24151). The sequence shows significant homology with the
Db	21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40	CC	conglutin family of seed storage proteins of other legumes. The allergen
QY	122 CCCTGCGAGCAACATCTCATGCAAGAGATCCAAACGTCGAGGATTCATATGAACGGGAC 181	CC	is recognised by serum IgE from a large proportion of individuals with
Db	41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60	CC	peanut hypersensitivity. Ara hII and Ara hI (see AAW24149-50) can be used
QY	182 CCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241	CC	to raise monoclonal antibodies which are used in a specific two-site MAB
Db	61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80	CC	ELISA for the detection of Ara hI or Ara hII (claimed). IgE- binding Ara
QY	242 TCCTCTCAGCAACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301	CC	hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect
Db	81 SerSerGlnHisGlnArgCysAsnGlnLeuAsnGluPheGluAsnAsnGlnArg 100	CC	against allergic reactions to peanut allergens, e.g. anaphylactic shock.
QY	302 TGCATGTGCGAGGATTCGAACACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGGAGG 361	CC	(Updated on 17-OCT-2003 to standardise OS field)
Db	101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120	XX	Sequence 157 AA;
QY	362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCTCAACAGTCGGCCCTTAGG 421		Alignment Scores:
Db	121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140		Pred. No.: 8.2e-90 Length: 157
QY	422 GCACCCAGCGTTGGCACTTGGACGTGAAAGTGGGGCGGACAGACATAC 472		Score: 842.00 Matches: 157
Db	141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157		Percent Similarity: 100.00% Conservative: 0
RESULT 2			Best Local Similarity: 100.00% Mismatches: 0
AAW24164			Query Match: 66.56% Indels: 0
ID	AAW24164 standard; protein; 157 AA.		DB: 2 Gaps: 0
XX			
AC	AAW24164;		
XX			
DT	17-OCT-2003 (revised)		
DT	29-DEC-1997 (first entry)		
XX			
DE	Peanut allergen Ara hII.		
XX			
KW	Peanut; seed storage protein; allergen; allergy; hypersensitivity;		

Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
 Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAenLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCACAGCTTCGCACTGGAGCTCGAAGTGGCGGCAGACACATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 3
 AAY15245
 ID AAY15245 standard; protein; 157 AA.
 AC AAY15245;
 XX
 DT 17-OCT-2003 (revised)
 DT 09-NOV-1999 (first entry)
 XX
 DE Peanut allergen, Ara h 2, amino acid sequence.
 XX
 KW allergy; immune response; transgenic; allergen; epitope;
 KW immunoglobulin E; Ig E; binding site; peanut.
 XX
 OS Arachis hypogaea.
 XX
 PN WO9938978-A1.
 XX
 PD 05-AUG-1999.
 XX
 PF 29-JAN-1999; 99WO-US002031.
 XX
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 27-AUG-1998; 98US-00141220.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
 PA (SOSI/) SOSIN H.
 XX
 PI Sosin H, Bannon GA, Burks AW, Sampson HA;
 XX
 DR WPI; 1998-479189/40.
 DR N-PSDE; AAZ06383.
 XX
 PT Modified allergen with reduced IgE binding, useful for treating e.g.
 PT allergies.
 XX
 PS Disclosure; Page 38; 46pp; English.
 XX
 CC This is the amino acid sequence of the Ara h 2 protein from Arachis
 CC hypogaea. The Ara h 2 protein has 10 IgE (Immunoglobulin E) binding
 CC epitopes, three of which are immunodominant (AAY15272, AAY15275, and
 CC AAY15276). By modifying the IgE binding sites the ability of the allergen
 CC to provoke an immune response is downregulated. The epitopes of the IgE
 CC binding sites can therefore be modified in genetically engineered plants
 CC and animals to elicit less of an allergic response. (Updated on 17-OCT-
 CC 2003 to standardise OS field)
 XX
 SQ Sequence 157 AA;

Alignment Scores:
 Pred. No.: 8.2e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 66.56% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAY15245 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTTCCTCCTCGCTGCCACCGCATCTCGAGGCAG 61
 Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20
 Qy 62 CAGTGGGAACCTCCAGGAGACAGAAAGATGCCAGAGCCAGCTCCGAGAGGGCGAACTCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAenLeuArg 40
 Qy 122 CCTCGAGCAACATCTCATGTCAGAGATCCAACTGACCGTGCAGAGATTTCATATGACCGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTTGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
 Qy 242 TCCTCTCAGCAACCAAGAGAGAGGTGTTCATAGCTGAACGAGTTTGAGAAACAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysCysAenGluLeuAenGluPheGluAenAenGlnArg 100
 Qy 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCCAGAGCCATAGTTTCAGGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCTCAACAGTGGCGGCAGACATAC 421
 Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAenLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCACAGCGTTGCGACTTGGACCTCGAAGTGGCGGCAGAGACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 4
 AAU04707
 ID AAU04707 standard; protein; 157 AA.
 AC AAU04707;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Anaphylactic antigen Ara h 2.
 XX
 KW Ara h 2; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
 KW allergy; mast cell; basophil; mouse.
 XX
 OS Mus sp.
 XX
 PN WO200140264-A2.
 XX
 PD 07-JUN-2001.
 XX
 PF 06-DEC-2000; 2000WO-US033124.
 XX
 PR 06-DEC-1999; 99US-00455294.
 PR 23-JUN-2000; 2000US-021376SP.
 PR 27-SEP-2000; 2000US-0235797P.
 XX
 PA (PANA-) PANACEA PHARM LLC.
 PA (UYAR-) UNIV ARKANSAS.
 PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
 XX
 PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
 XX
 DR WPI; 2001-381378/40.
 XX
 PT Antigenic fragments useful for reducing anaphylactic risk and reducing
 PT the severity and/or number of allergic symptoms in individuals sensitive
 PT to antigens, have reduced ability to bind Immunoglobulin E.
 XX
 PS Claim 7; Fig 10; 100pp; English.
 CC
 CC The sequence represents the amino acid sequence of anaphylactic antigen
 CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design

CC antigenic peptides having a reduced ability to bind immunoglobulin E
CC (IgE) as compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8.2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x AAU04707 (1-157)

QY 2 CTCACCATAGTAGGCGCTTCCTCTCTGCTGCCACGCGATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAAGAGACAGAAATGCCAGGCGATCCAGAGGGCGCAACTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGTGGAGCAATCTCTATGACAGATCCACGTCGACGAGGATTCATATGAACGGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLeuValIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGATACAGCCCTAGTCATATGATCGAGAGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGACCAAGAGAGGTGTTCATAGCTGAACGAGTTTGAGAACCAACGAAG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCCAGGCGATTCGCAACAGATCATGAGAACCCAGAGCGATAGTTGACAGGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTTCAAGAGGAGGCTCAGGAACCTTGCTCAACAGTGCGGCTTAGG 421
DB 121 GlnGlnGluGlnGlnPheGlyArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGCGTTGCGAGCTTGGACGTCGAAAGTGGGGCGGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 5
ABU52464
ID ABU52464 standard; protein; 157 AA.

XX
AC ABU52464;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h2 protein.
XX
KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US0009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI; 2003-018765/01.
DR N-PSDB; ABX70606.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Claim 27; Fig 41; 300pp; English.
XX

CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen (e.g. Ara h1, h2 or h3)
XX

SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8.2e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x ABU52464 (1-157)

QY 2 CTCACCATAGTAGGCGCTTCCTCTCTGCTGCCACGCGATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAAGAGACAGAAATGCCAGGCGATTCATATGAACGGGAC 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCGTACAGCCCTAGTCAGGATCCGATACAGCCCTAGTCATATGATCGAGAGCGCTGGA 181
DB 41 ProCysGluGlnHisLeuMetGlnLeuValIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGATACAGCCCTAGTCATATGATCGAGAGCGCTGGA 241
DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80

QY 242 TCCTCTGAGCACCAGAGAGGTGTCATGAGCTGAACGAGTTTGAGAACACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
 QY 302 TGCATGTGCGAGGCAATTCACACAGATCATGAGAACACAGAGCGATAGTTTCAGGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CAACAGGAGCAACAGTTCAACAGGAGGCTCAGGAATTCCTCAACAGTGGCGCTTAGG 421
 Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACACAGCGTTTCGCACTTCGACGTGCAAGTGGCGGCGCAGACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 6
 ID ADG27518 standard; protein; 157 AA.
 AC ADG27518;
 XX 26-FEB-2004 (first entry)
 DT XX
 DE Peanut allergen Ara h2.
 XX Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycine A2Bla; Jug n1;
 KW antiallergic; vulnerary; anaphylactic food allergen; IgE; allergy; wound.
 XX
 OS *Arachis hypogaea*.
 XX US2003202980-A1.
 PN 30-OCT-2003.
 XX 18-MAR-2002; 2002US-00100303.
 XX 29-DEC-1995; 95US-0009455P.
 PR 23-SEP-1996; 96US-00717933.
 PR 31-JAN-1998; 98US-0073283P.
 PR 13-FEB-1998; 98US-0074590P.
 PR 13-FEB-1998; 98US-0074624P.
 PR 13-FEB-1998; 98US-0074633P.
 PR 29-JUN-1998; 98US-00106872.
 PR 27-AUG-1998; 98US-00141220.
 PR 13-NOV-1998; 98US-00191593.
 PR 29-JAN-1999; 99US-00240557.
 PR 29-JAN-1999; 99US-00241101.
 PR 11-FEB-1999; 99US-00248673.
 PR 11-FEB-1999; 99US-00248674.
 PR 02-MAR-1999; 99US-0122450P.
 PR 02-MAR-1999; 99US-0122452P.
 PR 02-MAR-1999; 99US-0122560P.
 PR 02-MAR-1999; 99US-0122565P.
 PR 02-MAR-1999; 99US-0122566P.
 PR 11-MAR-1999; 99US-00267719.
 PR 28-JAN-2000; 2000US-00494096.
 PR 16-MAR-2001; 2001US-0276822P.
 XX (CAPL/) CAPLAN M. J.
 PA (SOSI/) SOSIN H. B.
 PA (SAMP/) SAMPSON H.
 PA (BANN/) BANNON G. A.
 PA (BURK/) BURKS A. W.
 PA (COCK/) COCKRELL G.
 PA (COMP/) COMPADRE C. M.
 PA (CONN/) CONNAUGHTON C.
 PA (HELM/) HELM R. M.
 PA (KING/) KING N. E.
 PA (KOPP/) KOPPER R. A.
 PA (MALE/) MALEKI S. J.
 PA (RABJ/) RABJOHN P. A.

PA (SHIN/) SHIN D. S.
 PA (STAN/) STANLEY J. S.
 PI Caplan MJ, Sosin HB, Sampson H, Bannan GA, Burks AW, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabinjohn PA, Shin DS, Stanley JS;
 XX WPI; 2003-875632/81.
 DR N-PSDB; ADG27517.
 XX New modified anaphylactic food allergen comprising a cysteine residue
 PT which has been modified so that it cannot participate in the disulfide
 PT bond, useful for treating allergic reactions or wounds.
 XX Claim 27; SEQ ID NO 63; 194pp; English.
 XX The invention relates to a modified anaphylactic food allergen whose
 CC amino acid sequence is substantially identical to that of a natural
 CC anaphylactic food allergen. The natural anaphylactic food allergen
 CC includes at least one cysteine residue that participates in a disulfide
 CC bond when the natural anaphylactic food allergen is in its native
 CC conformation, except that the cysteine residue has been modified so that
 CC it cannot participate in the disulfide bond. Also included are a method
 CC of making a modified anaphylactic food allergen, a nucleotide molecule
 CC encoding a modified anaphylactic food allergen defined above, a
 CC nucleotide molecule for causing a site specific mutation in a gene
 CC encoding a natural anaphylactic food allergen, a transgenic plant or
 CC animal expressing a modified anaphylactic food allergen defined above, a
 CC method of treating an individual by reducing the clinical response to a
 CC natural anaphylactic food allergen by administering a modified
 CC anaphylactic food allergen and an isolated fragment of peanut allergen
 CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
 CC ADG27465. About 10-17% of the amino acids have been modified in at least
 CC one IgE epitope or all the IgE epitopes recognised when the natural
 CC anaphylactic food allergen is contacted with serum IgE from individual(s)
 CC allergic to the natural anaphylactic food allergen. The invention
 CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
 CC cDNAs), Soybean Glycinin A2Bla and IgE-binding epitopes of the English
 CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
 CC used for treating allergic reactions or wounds. The present sequence
 CC represents a Peanut allergen of the invention (or its fragment).
 XX
 SQ Sequence 157 AA;
 Alignment Scores:
 Pred. No.: 8,2e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 66.56% Indels: 0
 DB: 7 Gaps: 0
 US-10-728-323-2 (1-717) x ADG27518 (1-157)
 QY 2 CTCACCATACTAGTACGCTCGCCCTTCCTCTCGCTGCCACGCGATCTGGAGGAG 61
 Db 1 LeuthrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAAGTCCCAAGGAGACAGAGTGCAGAGCGAGCTCGAGGGGGGAACCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCCTGCGAGCAACATCTCATGCGAGAAGATCCAACTGACGAGGATTCATATGAACGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CGGTACAGCCCTAGTACGAGTCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
 QY 242 TCCTCTCAGCAACCAAGAGAGGTGTCGAATGACCTGAACGAGTTTCAGAACCAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100

QY	302	TGCATGTGCGAGGCATTGCAACAGATCATGAGAACACAGACGATAGGTTGCGAGGGAGG	361
Db	101	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	120
QY	362	CAACAGGAGCAACAGTTCAAGAGGGAGGCTCAGGAACCTTGCTCAACAGTCGCGCCTTAGG	421
Db	121	GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	140
QY	422	GCACCCAGCGTTGCGACTTGGACGTGCGAAGTGGCGGCAGACAGACATAC	472
Db	141	AlaProGlnArgCysAspLeuAepValGluSerGlyArgAspArgTyr	157
RESULT 7			
AA	B82383	standard; protein; 207 AA.	
XX	AC	AAB82383;	
XX	DT	11-SEP-2003 (revised)	
XX	DT	23-JUL-2001 (first entry)	
DE		Peanut allergen Ara h2 gene product.	
XX	KW	Peanut; allergen; Ara h2; transgenic plant; allergy.	
XX	OS	Arachis hypogaea.	
XX	XX	Location/Qualifiers	
FT	Key	1. .21	
FT	Peptide	/label= Signal_peptide	
FT	Protein	22. .207	
FT	Protein	/label= Mature_protein	
XX	XX	WO200136621-A2.	
XX	PD	25-MAY-2001.	
XX	PF	20-NOV-2000; 2000WO-US031657.	
XX	PR	19-NOV-1999; 99US-0167255P.	
XX	XX	(UYAL-) UNIV ALABAMA A & M.	
XX	PA	Dodo HW, Arntzen CJ, Konan KN, Viquez OM;	
XX	PI	WPI; 2001-355630/37.	
XX	DR	N-PSDB; AAF90336, AAF90337.	
XX	PT	Producing transgenic peanut plants that produce allergen-free seeds, useful in non-allergenic foods, by antisense or sense co-suppression of allergen-encoding genes.	
XX	PS	Example 1; Fig 2; 72pp; English.	
XX	CC	The present sequence is that of the peanut allergenic protein (AP) encoded by the Ara h2 gene (see AAF90336). The invention relates to a method for producing a peanut plant having reduced, or undetectable, AP content in its seed. A peanut plant cell is transformed with a DNA construct containing an antisense AP gene and/or sense AP gene, or their fragments, regenerated to plants, and fertile transgenic plants that produce seeds with reduced AP content are identified. The AP sense or antisense gene may comprise at least a portion of the Ara h2 gene sequence. The seeds are useful for preparation of allergen-free foods. Recombinant AP may be produced and used to produce antibodies useful for detecting AP in foods, and for treatment or prevention of peanut allergy. (Updated on 11-SEP-2003 to standardise OS field)	
XX	SQ	Sequence 207 AA;	
Alignment Scores:			
Pred. No.:	2.39e-89	Length:	207
Score:	838.50	Matches:	162

Percent Similarity:		90.27%	Conservative:	5
Best Local Similarity:		87.57%	Mismatches:	17
Query Match:		66.28%	Indels:	1
DB:		4	Gaps:	1
US-10-728-323-2 (1-717) x AAB82383 (1-207)				
QY	2	CTCACCATCTAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGCAG	61	
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln	23	
QY	62	CAGTGGAACTCCAGAGAGACAGATGCCAGAGCCAGCTCGAGAGCGCAACCTGAGG	121	
Db	24	GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	43	
QY	122	CCCTGCGAGCAACATCTCATGCAAGATCCCAACGTGACGAGGATTCATATGAACGGAC	181	
Db	44	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	63	
QY	182	CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATATCGAGAGCGCTGGA	241	
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	83	
QY	242	TCCTCTCAGCACCAAGAGAGGTGTCGAATGAGCTGAACGAGTTTGAGAACCAACG	301	
Db	84	SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	103	
QY	302	TGCATGTGCGAGGCATTGCAACAGATCATGAGAACACAGACGATAGGTTGCGAGGAGG	361	
Db	104	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	123	
QY	362	CAACAGAGCAACAGTTCAAGAGGAGCTTCAGGAACCTTGCTCAACAGTCGCGCCTTAGG	421	
Db	124	GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	143	
QY	422	GCACCCAGCGTTGCGACTTGGACGTGCGAAGTGGCGGCAGACACAGATAC	481	
Db	144	AlaProGlnArgCysAspLeuAepValGluSerGlyArgProArgIleProPro	163	
QY	482	ATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGC--TATTATCTA	538	
Db	164	IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer	183	
QY	539	TGTTATGTTTAGTT 553		
Db	184	AlaMetCysLeuLeu 188		
RESULT 8				
AA	Y40973	standard; protein; 156 AA.		
XX	AC	AAAY40973;		
XX	DT	17-OCT-2003 (revised)		
XX	DT	06-DEC-1999 (first entry)		
XX	DE	Ara h 2 protein fragment.		
XX	KW	Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;		
XX	KW	allergic reaction; Ara h 2.		
XX	OS	Arachis hypogaea.		
XX	PN	WO9945961-A1.		
XX	XX	16-SEP-1999.		
XX	PF	12-MAR-1999; 99WO-US005494.		
XX	PR	12-MAR-1998; 98US-0077763P.		
XX	PR	11-MAR-1999; 99US-00077763.		
XX	XX	(UYAR-) UNIV ARKANSAS.		

XX Burks W, Helm RM, Cockrell G, Bannan GA, Stanley JS, Shin DS;
 XX Sampson H, Compadre CW, Huang SK, Maleki SJ, Kopper RA;
 XX WPI; 1999-551218/46.
 XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
 XX animal from allergic reaction.
 XX Disclosure; Page 104; 193pp; English.
 XX The invention provides a tertiary structure for the peanut allergen Ara H
 CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
 CC epitopes. The invention also provides an isolated recombinant peanut
 CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
 CC allergen Ara h 3. Molecules of the invention are used to protect a host
 CC animal from allergic reaction, particularly using a modified allergen
 CC which is less reactive with IgE. The invention may also be used to ensure
 CC that the allergen is not introduced into genetically modified food. The
 CC present sequence represents an Ara h 2 protein fragment. (Updated on 17-
 CC OCT-2003 to standardise OS field)
 XX
 SQ Sequence 156 AA;
 Alignment Scores:
 Pred. No.: 2.13e-88 Length: 156
 Score: 830.00 Matches: 155
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 65.61% Indels: 0
 DB: 2 Gaps: 0
 US-10-728-323-2 (1-717) x AAY40973 (1-156)
 QY 2 CTCACCTACTAGTAGCCCTCGCCCTTTCTCTCTCTGCGCCAGCATCTGCGAGGAG 61
 DB 1 LeuThrIleuValaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 20
 QY 62 CAGTGGGAATCCAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGCGGCAACCTGAGG 121
 DB 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 QY 122 CCCTGCGAGCAACATCTTCAGAGAGATCCAGCTGACAGAGATTCATATGACGGGAC 181
 DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTTGA 241
 DB 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGAGCTGAGTGAACGAGTTTGAGAACCAAGG 301
 DB 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 QY 302 TGCATGTCGAGGAGGATTGCAACAGATCATGAGAACAGAGCCAGATAGTTGCGAGGAGG 361
 DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 QY 362 CACAGGAGCAACAGTTTCAGAGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCTTAGG 421
 DB 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 QY 422 GCACCACAGGCTTGCAGCTTGGAGCTGCAAGTGGCGGCGGACAGAC 466
 DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAsp 155
 RESULT 9
 ADO38314
 ID ADO38314 standard; peptide; 156 AA.
 XX
 AC ADO38314;
 XX
 DT 15-JUL-2004 (first entry)

XX Ara h 2 MHC-class II-presented epitope #1.
 XX Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
 XX Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
 XX Immunosuppressive; Antidiabetic; Antithyroid; Antiasthmatic;
 XX Antiallergic; Cytostatic; Antipsoriatic; Gene Therapy; Vaccine;
 XX MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
 XX influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
 XX rickettsia; rheumatoid arthritis; multiple sclerosis;
 XX lupus erythematosus; diabetes mellitus; myasthenia gravis;
 XX autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
 XX allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
 XX adenoma; peanut; Ara h 2.
 XX Arachis hypogaea.
 OS US2004058881-A1.
 PN 25-MAR-2004.
 PD 24-SEP-2002; 2002US-00253286.
 PF 24-SEP-2002; 2002US-00253286.
 PR (ANTI-) ANTIGEN EXPRESS INC.
 XX Humphreys RE, Xu M;
 XX WPI; 2004-294259/27.
 XX New non-naturally occurring protein or polypeptide modified by
 PT recombinant DNA techniques, useful for treating multiple sclerosis,
 PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
 PT colitis, cancer or psoriasis.
 XX Example 2; Page 18; 90pp; English.
 PS
 XX The invention relates to a non-naturally occurring protein or polypeptide
 CC (I) modified by recombinant DNA techniques comprising: a C-terminal
 CC element comprising an MHC Class II-presented epitope; an N-terminal
 CC element comprising an Ii-key motif; and an intervening element comprising
 CC a sequence of 4-11 amino acid residues where the modification by
 CC recombinant DNA techniques taking place within elements (b) and (c). Also
 CC described are methods for: suppressing or enhancing an immune response
 CC directed toward an MHC (major histocompatibility complex) Class II-
 CC presented epitope of interest. Suppressing an immune response directed
 CC toward an MHC Class II-presented epitope of interest comprises: providing
 CC a nucleic acid sequence encoding the MHC Class II-presented epitope of
 CC interest, the nucleic acid sequence encoding an Ii-key motif located 4-11
 CC amino acids upstream from the N-terminal residue of the MHC Class II-
 CC presented epitope of interest; and modifying the Ii-key motif to decrease
 CC its conformance to the archetypal Ii-key regulatory motif. Enhancing an
 CC immune response directed toward an MHC Class II-presented epitope of
 CC interest comprises: providing a nucleic acid sequence encoding the MHC
 CC Class II-presented epitope of interest, the nucleic acid sequence lacking
 CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal
 CC residue of the MHC Class II-presented epitope of interest; and modifying
 CC the nucleic acid sequence to introduce an Ii-key motif appropriately
 CC spaced from the MHC Class II-presented epitope. The protein or
 CC polypeptide of interest corresponds to a protein or polypeptide encoded
 CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,
 CC preferably vaccinia virus. The non-naturally occurring protein or
 CC polypeptide (I) modified by recombinant DNA techniques is useful for
 CC treating infectious diseases caused or associated with infection by a
 CC bacterium, virus, parasite, fungus, rickettsia or other infectious
 CC agents. It is also useful for treating rheumatoid arthritis, multiple
 CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,
 CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,
 CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or
 CC adenomas. The present sequence represents the amino acid sequence of a
 CC peanut allergen Ara h 2 MHC class II-presented epitope used in the
 CC invention.

Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 11

ID ADM12096 standard; protein; 156 AA.

XX AC ADM12096;

DT 20-MAY-2004 (first entry)

XX DE Arachis hypogaea 2 (Ara h2) protein.

XX KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-key;
KW MHC Class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.

OS Arachis hypogaea.

XX PN US2003235594-A1.

XX PD 25-DEC-2003.

XX PF 17-SEP-2002; 2002US-00245871.

XX PR 14-SEP-1999; 99US-00396813.

XX PR 17-JUL-2002; 2002US-00197000.

XX PA (ANTI-) ANTIGEN EXPRESS INC.

XX PI Humphreys R, Xu M;

XX DR WPI; 2004-070554/07.

XX PT Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.

XX PS Example 2; Page 19; 87pp; English.

XX CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian Ii-key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC represents a mammalian Ii key related protein of the invention.

XX SQ Sequence 156 AA;

Alignment Scores:

Pred. No.: 2,83e-86 Length: 156
Score: 812.00 Matches: 152
Percent Similarity: 99.35% Conservatives: 0
Best Local Similarity: 99.35% Mismatches: 1
Query Match: 64.19% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-2 (1-717) x ADM12096 (1-156)

QY 2 CTCACACTACTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCCGCATCTCGAGGCAG 61
DB |||||||
4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGAACTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCCGAGAGGGCGAACCTGAGG 121
DB |||||||
24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGGAGCAACATCTCATGCAAGATCCAAAGATCCAAAGTGAACGAGATTTCATATGAACGGGAC 181
DB |||||||
44 ProCysGluGlyHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCTCAGAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
DB |||||||
64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTTGTCATAGCTGAACAGTTTTCAGAAACAACCAAGG 301
DB |||||||
84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGAGAACCCAGAGCGATAGTTTGCAGGGGAGG 361
DB |||||||
104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CACAGGAGCCACAGTTCAGAGGGAGCTCAGGAACCTTCCTCAACAGTCCGCTTAGG 421
DB |||||||
124 GlnGlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACACAGCGTTGCGACTTGGACGTGGAAGTGGGGCGC 460
DB |||||||
144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 12
ABU52575
ID ABU52575 standard; protein; 157 AA.
XX AC ABU52575;
XX DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h2 mutant 4 (E35A/P41A/D60A/D67A).
XX KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
XX KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
XX KW wound healing.
XX OS Homo sapiens.
XX PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX PA (PANA-) PANACEA PHARM.
XX PI Caplan M, Sosin H, Sampson H, Bannan CA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
XX PT New modified anaphylactic food allergen, useful for preventing or
XX treating allergic reactions associated with e.g. anaphylactic allergens.
XX PS Example 13; Page; 300pp; English.
XX CC The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been

CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IGE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IGE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 2,84e-86 Length: 157
Score: 812.00 Matches: 153
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 4
Query Match: 64.19% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABUS2575 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAGAGACAGAGATGCGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAATCTCATGCAAGATCCAAACGTGACGAGGATTTCATATGAACGGGAC 181
DB 41 AlaCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAla 60
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241
DB 61 ProTyr-SerProSerGlnAlaProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTCATGACTGACCTGACAGTTTCAGAACACCAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGATGTGCGAGGCATTGCAACAGATCATGAGAACCCAGAGCGATAGGTTGCAAGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG 421
DB 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGCGTTCGACCTGGACGTGCAAGTGGCGGCGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 13
ABUS2577

ID ABUS2577 standard; protein; 157 AA.
XX
AC ABUS2577;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h2 mut. (W22A/E35A/P41A/D53A/D60A/D67A/R120A/L130A/L147A).
XX
KW Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;

KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
KW wound healing.
XX
OS Homo sapiens.
XX
PN WO200274250-A2.
XX
XX 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CW, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 13; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IGE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a food
XX allergen, mutated to alter its IGE binding characteristics. Note: The
XX present sequence is not shown in the specification but was created by the
XX indexer using information provided in the specification
XX
SQ Sequence 157 AA;

Alignment Scores:
Pred. No.: 8,63e-82 Length: 157
Score: 774.00 Matches: 148
Percent Similarity: 94.27% Conservative: 0
Best Local Similarity: 94.27% Mismatches: 9
Query Match: 61.19% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABUS2577 (1-157)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGCATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAGAGACAGAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
DB 21 GlnAlaGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArgAlaAsnLeuArg 40
QY 122 CCTGCGAGCAATCTCATGCAAGATCCAAACGTGACGAGGATTTCATATGAACGGGAC 181
DB 41 AlaCysGluGlnHisLeuMetGlnLysIleGlnArgAlaGluAspSerTyrGluArgAla 60
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241

Db 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgAGrGlyAlaGly 80
Qy 242 TCCTCTCAGCACCAAGAGAGGTGTTCACATGAGCTGACGAGTTTGAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAenGluPheGluAenAsnGlnArg 100
Qy 302 TGCATGTGCGAGGCAATTCACACAGATCATGAGNACACAGAGCGATAGTTGCAGGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArgLeuGlnGlyAla 120
Qy 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTGCTCAACAGTGGCGGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluAlaAargAsnLeuProGlnGlnCysGlyLeuArg 140
Qy 422 GCACCAAGCGTTCCGACTTGACGTGCAAGAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 14
ID AAY40968 standard; protein; 166 AA.
AC AAY40968;
XX
XX
DT 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX
XX
XX Recombinant Ara h 2 protein sequence.
XX
XX Peanut; allergen; Ara h 1; IGE; immunoglobulin E; epitope; Ara h 3;
XX allergic reaction; Ara h 2.
XX
XX Arachis hypogaea.
XX
XX WO9945961-A1.
XX
XX 16-SEP-1999.
XX
XX 12-MAR-1999; 99WO-US005494.
XX
XX 12-MAR-1998; 98US-0077763P.
PR 11-MAR-1999; 99US-00077763.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX
XX WPI; 1999-551218/46.
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
XX
XX Disclosure; Page 81; 193pp; English.
XX
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IGE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IGE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents a recombinant Ara h 2 protein. (Updated on 17
CC -OCT-2003 to standardise OS field)
XX
XX Sequence 166 AA;
SQ
Alignment Scores:
Pred. No.: 2e-81 Length: 166
Score: 771.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-2 (1-717) x AAY40968 (1-166)
Qy 50 TCTGCGAGGCGAGCAGTGGGAACCTCCAGGAGAGACAGAGATGCCAGAGCCAGCTCGAGAGG 109
Db 15 SerAlaAargGlnGlnTrpGluLeuGlnGlnIleMetGluAenGlnSerGlnLeuGluArg 34
Qy 110 GCCAACTCAGGCGCTCGAGGCAACATCTCATCAGAGAAGATCCAACTGACGAGGATTCA 169
Db 35 AlaAsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer 54
Qy 170 TATGAACGGGACCCCGTACAGCCCTAGTACAGATCCGTTACAGCCCTAGTCCATATGATCGG 229
Db 55 TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 74
Qy 230 AGAGGCGCTCGATCCTCTCAGCAGCAGGAGAGAGGTGTTCATGAGCTGAACGAGTTTGG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAenGluPheGlu 94
Qy 290 AACAAACCAAGGTGCATGTGCGAGGCAATTCGCAACAGATCATGAGAGAACAGAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArg 114
Qy 350 TTCAGGGGAGGCAACAGAGGCAACAGTTCAAGAGGAGCTCAGGAACTTGCCTCAACAG 409
Db 115 LeuGlnGlyArgGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln 134
Qy 410 TGGCGGCTTAGGCGCACAGCGTTCGAGCTTCGAGCTGAAAGTGGCGGCGAGACAGAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 154
Qy 470 TAC 472
Db 155 Tyr 155
RESULT 15
ABU52482
ID ABU52482 standard; protein; 166 AA.
XX
XX AC ABU52482;
XX
XX 10-MAR-2003 (first entry)
XX
XX Peanut Ara h2 with N-terminal a T7 tag and a C-terminal His tag.
XX
XX Peanut; allergy; Ara h1; Ara h2; Ara h3; IGE binding site;
XX anaphylactic food allergen; anti-allergenic; vaccine; wound healing.
XX
XX Arachis hypogaea.
OS Synthetic.
OS Enterobacteria phage T7.
XX
XX WO200274250-A2.
XX
XX 26-SEP-2002.
XX
XX 18-MAR-2002; 2002WO-US0009108.
XX
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
XX
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NS, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT

PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX Example 13; Fig 51; 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen, Ara h2, with an N-terminal 17 tag and a C-terminal His
CC tag

XX Sequence 166 AA;

Alignment Scores: 2e-81 Length: 166
Pred. No.: 771.00 Matches: 141
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-1717) x ABUS2482 (1-166)

Qy	50	TC	TCCGAGGAGCAGTGGGAATCC	AAGACAGAGATGCCAGAGCCAGCTCGAGAGG	109
Db	15	Ser	AlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArg		34
Qy	110	GC	AACTGAGGCCCTCGAGCAACATCT	CATCAGAAATCCAAAGTGACGAGGATTCA	169
Db	35	Ala	AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer		54
Qy	170	TAT	GAACGGACCGTACACCCCTAGT	CAGATCCGTCAGCCCTAGTCCATATGATCGG	229
Db	55	Tyr	GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg		74
Qy	230	AG	AGCGCTCGATCCTCTCAGCA	CCAGAGAGGTGTGCAATCAGCTGAACGAGTTTGAG	289
Db	75	Arg	GlyAlaGlySerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGlu		94
Qy	290	AAC	ACCAAGTGCATGTGCGAGGCNTT	GCACAGATCATGAGAACCGAGCGATAGG	349
Db	95	Asn	AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg		114
Qy	350	TT	CAGGGAGGCAACAGGAGCAACAGTT	CAAGAGGAGCTCAGGAACCTTGCCCTCAACAG	409
Db	115	Leu	GlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln		134
Qy	410	TG	CGGCTTAGGCACACAGCGTTGCGA	CTTGAGCTCGAAAGTGGCGGACAGACAGA	469
Db	135	Cys	GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg		154
Qy	470	TAC	472		
Db	155	Tyr	155		

Search completed: August 24, 2005, 09:55:57
Job time : 43.6071 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 07:15:51 ; Search time 8.97718 Seconds
(without alignments)
15369.504 Million cell updates/sec

Title: us-10-728-323-2

Perfect score: 1265

Sequence: 1 gctcacatagtagcc.....taaaagatcatgttttgg 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10728323/runat 23082005 124355 29222/app_query.fasta_1.4757
-DB=PIR_79 -QMT=fastan -SUFFIX=n2p.tpr -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40 cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1_1_76 @runat 23082005 124355 29222 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	278	22.0	153	2 A33090	conglutin delta pr
2	243	19.2	158	1 T05710	2S albumin precurs
3	187.5	14.8	80	2 B23617	conglutin delta-2
4	184.5	14.6	323	2 S38887	2S albumin - commo
5	173.5	13.7	141	2 T10257	2S albumin precurs
6	154	12.2	154	2 S14947	2S albumin - Brazi
7	151	11.9	295	2 S01062	2S seed storage pr
8	147	11.6	186	2 A29802	napin precursor (g
9	146.5	11.6	164	1 NWU01	2S albumin I precu
10	140.5	11.1	146	2 S14946	2S seed storage pr
11	139	11.0	155	2 JC5379	mablin II precu
12	139	11.0	258	1 RZCS	2S seed storage pr
13	132.5	10.5	165	2 T09252	seed storage prote
14	129	10.2	170	2 T08012	2S seed storage pr

15	128	10.1	172	2 S18871	2S-like storage pr
16	126.5	10.0	70	2 A59346	seed storage prote
17	125	9.9	162	2 T08013	2S seed storage pr
18	124.5	9.8	162	2 A92559	albumin 4 - easter
19	122.5	9.7	162	2 T08010	2S seed storage pr
20	121	9.6	100	2 S48180	mablin IV - Yun
21	121	9.6	174	2 S80425	napin B3 precursor
22	119.5	9.4	180	2 S52025	napin (clones BNMN
23	118.5	9.4	106	2 S26636	napin n1b - rape
24	118.5	9.4	110	2 S20350	napin n1a - rape
25	118	9.3	104	2 S48176	mablin I-1 - Yun
26	117	9.2	104	2 S48178	mablin III - Yun
27	115	9.1	152	2 S80427	napin AHI precursor
28	111.5	8.8	173	2 T08011	2S seed storage pr
29	108.5	8.6	178	2 S07828	napin B - rape
30	108.5	8.6	178	2 S25127	2S storage protein
31	107	8.5	139	2 T09850	albumin 2S storage
32	107	8.5	139	2 T09878	albumin 2S storage
33	106.5	8.4	178	1 NMRP2	napin 2 precursor
34	106.5	8.4	178	2 A25997	napin precursor (n
35	105.5	8.3	164	1 NWMU3	napin (clone BngNA
36	105.5	8.3	180	2 S10018	tumor suppressor p
37	105.5	8.3	1099	2 A56155	liver stage anticp
38	105.5	8.3	1909	2 A45592	hypothetical prote
39	104.5	8.3	1263	2 T15496	allergen Sin a I (
40	104	8.2	145	2 S65479	allergen Sin a I (
41	103.5	8.2	166	1 NWMU4	2S albumin 4 precu
42	103.5	8.2	178	2 S25130	2S storage protein
43	102.5	8.1	178	2 S25134	2S storage protein
44	101.5	8.0	178	2 S25137	2S storage protein
45	101	8.0	145	2 PC1246	Sin a I allergen 1

ALIGNMENTS

RESULT 1

A33090 conglutin delta precursor - narrow-leaved blue lupine

C/Species: Lupinus angustifolius (narrow-leaved blue lupine)
C/Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 09-Jul-2004

C/Accession: S12404; A33090

R/Gayler, K.R.; Kolivas, S.; MacFarlane, A.J.; Lilley, G.G.; Baldi, M.; Blagrove, R.J.;

Plant Mol. Biol. 15, 879-893, 1990

A/Title: Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta,

A/Reference number: S12404; MWID:91355912; PMID:2103479

A/Accession: S12404

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-153 <GAY>

A/Cross-references: UNIPROT:Q99235; GB:X53523; NID:g19140; PIDN:CAA37598.1; PID:g19141

C/Superfamily: soybean 2S albumin

C/Keywords: seed

F/1-22/Domain: signal sequence #status predicted <SIG>

F/23-59/Product: conglutin small chain #status predicted <MAT1>

F/73-153/Product: conglutin large chain #status predicted <MAT2>

Alignment Scores:

Pred. No.:	3-84e-18	Length:	153
Score:	278.00	Matches:	62
Percent Similarity:	55.00%	Conservative:	26
Best Local Similarity:	38.75%	Mismatches:	46
Query Match:	21.98%	Indels:	26
DB:	2	Gaps:	5

US-10-728-323-2 (1-717) x A33090 (1-153)

QY	2	CTCACCATCTAGTAGCCCTC---GCCCTTTTCCTCTCGTCCGCCACGCATCTCGAGG 58
DB	4	LeuThrIleuLeuLeuValAlaLaLeuValLeuValValHisThrSerAlaPhe 23
QY	59	CAGCAGTGGGAATCCCAAGGACGACGAGATGCACGAGCCAGCTCGAGAGGGGGAACCTG 118
DB	24	Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnValAsnLeu 39

```
QY 119 AGCCCTGCAGCAACATCTCATGTCAGAAATCAACAGTCACGAGGATTCATATGAACGG 178
Db 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGlnGluGluAsp 59
QY 179 GACCCGTACAGC-----CCTAGTCAGGATCCG 205
Db 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerSerGlnGlu 79
QY 206 TACAGCCCTAGTCCATGATGCGAGAGGGCTGATCTCTCAGCACCAAGAGAGGTGT 265
Db 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCAAATGAGTGAACGAGTTTGAGAACAAACAAAGGTGCGATGTCAGGCGATTGCAACAG 325
Db 91 CysGluGlnLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGAGAACACAGAGCGATAGTTGTCAGGGAGGCAACAGAGCAACAGATTCACAGAG 385
Db 110 IleTyrGluSerGlnSerGlnGlnCysGluGlySerGlnGlnGlnGlnGlnGluGln 129
QY 386 GAGCTCAGGAACTTGCTCAACAGTTCGGCGCTTAGGGCCACCAGAGCGTTGCGACTTGGAC 445
Db 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValAsn 149

RESULT 2
T05710
2S albumin precursor - soybean
N;Alternate names: aspartic acid-rich peptide
C;Species: Glycine max (soybean)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: T05710; A28485
R;Wang, J.; Pichersky, E.
Plant Physiol. 114, 1567, 1997
A;Title: Nucleotide Sequence of S-Adenosyl-L-Methionine:Caffeic Acid 3-O-Methyltransferase
A;Reference number: Z15424
A;Accession: T05710
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-158 <W>N>
A;Cross-references: UNIPROT:P19594; EMBL:AF005030; NID:g2305019; PID:g2305020
A;Experimental source: cultivar Hodgson 78; cotyledon
R;Odani, S.; Koide, T.; Ono, T.
J. Biol. Chem. 262, 10502-10505, 1987
A;Title: Amino acid sequence of a soybean (Glycine max) seed polypeptide having a poly(L
A;Reference number: A28485; MUID:87280104; PMID:3611081
A;Accession: A28485
A;Molecule type: protein
A;Residues: 22-64 <ODA>
C;Superfamily: soybean 2S albumin
C;Keywords: glycoprotein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Product: 2S albumin small chain #status predicted <MAT>
F;54-56/Region: cell attachment (R-G-D) motif
F;39/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 7,86e-15 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x T05710 (1-158)

QY 5 ACCATAGTAGCCCTCGCCCTTTTCTCTCGCTGCCACGCATCTCGAGGCAGCAG 64
Db 5 ThrIleLeuLysLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGAGACAGAAAGATGCCAGAGCGAGCTCGAGAGGGGGAACCTGAGGCC 124
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlnGlyValAsnLeuThrPro 42
```

```
QY 125 TCGAGCAACATCTCATGTCAGAAATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GATCCTCTCTCAG-----CACCAGAGAGGCTTGCAATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGGAAACAACAAAGGTGCGATGTCAGGCGATTGCAACAGATCATGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 ACGATAGATTGCGAGGGAGGCAACAGGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTG 400
Db 122 SerGluGluGluGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTTCGGCGCTTAGGGCCACCAGAGCGTTGCGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
B23617
conglutin delta-2 large chain - narrow-leaved blue lupine
C;Species: Lupinus angustifolius (narrow-leaved blue lupine)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
C;Accession: B23617
R;Lilley, G.G.; Inglis, A.S.
FEBS Lett. 195, 235-241, 1986
A;Title: Amino acid sequence of conglutin delta, a sulfur-rich seed protein of Lupinus ar
A;Residues: 1-80 <LIL>
A;Cross-references: UNIPROT:P09931
C;Superfamily: soybean 2S albumin

Alignment Scores:
Pred. No.: 1.42e-09 Length: 80
Score: 187.50 Matches: 34
Percent Similarity: 73.13% Conservative: 15
Best Local Similarity: 50.75% Mismatches: 17
Query Match: 14.82% Indels: 1
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x B23617 (1-80)

QY 245 TCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGGTGC 304
Db 11 SerGluGluLeuAspGlnCysCysGluGlnLeuAsnGluLeu---AsnSerGlnArgCys 29
QY 305 ATGTGCGAGCGATTGCAACAGATCATGAGAACCAAGAGCGATAGGTTCGAGGGAGGCA 364
Db 30 GlnCysArgAlaLeuGlnGlnIleTyrGluSerGlnSerGluGlnCysGluGlyArgGln 49
QY 365 CAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTCGCGCCCTTAGGGCA 424
Db 50 GlnGluGlnGlnLeuGluGlyGluLeuGluLysLeuProArgIleCysGlyPheGlyPro 69
QY 425 CCACAGCGTTGCGACTTGGAC 445
Db 70 LeuArgArgCysAsnIleAsn 76

RESULT 4
S38887
2S albumin - common sunflower (fragment)
C;Species: Helianthus annuus (common sunflower)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S38887
```

Pred. No.:	2.07e-06	Length:	154
------------	----------	---------	-----

Score: 154.00 Matches: 45
Percent Similarity: 45.34%
Best Local Similarity: 27.95%
Query Match: 12.17%
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x S14947 (1-154)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTCGC----- 55
Db ValAlaAlaLeuAlaLeuValLeuGlyGlnAlaThrAlaPheAraThrThr 26
QY 56 -----AGCGACGAGTGGGAATCTCAAGGA-----GACAGAAGA 88
Db ValThrThrThrLeuGluGluGlnGluAsnProArgGlyArgSerGluGlnGln 46
QY 89 TGCAGAGCCAGCTCGAGAGG---GGAACTGAGGCCCTGCGAGCAACATCTCATGCGAG 145
Db CysArgGluGlnMetGluArgGlnGlnGlnLeuAsnHisCysArgMetTyrLeuArgGln 66
QY 146 AAGATCCAACTGACGAGGATTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCG 205
Db GlnMet-----GluGluSerProTyrGln----- 74
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCTCTCAGCAACGAAGAGAGTGT 265
Db AsnProArgProLeuArgAraG---GlyGluGluProHisLeuAspGluCys 90
QY 266 TGCAATGAGCTGAACAGATTGAGAACAAACAAAGGTGATGTCGAGGCAATGCAACAG 325
Db CysGluGlnLeuGluArgMetAsp-----GluMetCysArgCysGluGlyLeuArgMet 108
QY 326 ATCATGTGAGAACAGAGCGAT-----AGCTTCAGGGGGGCAACAGGAGCAACAGTTC 379
Db MetLeuArgArgGlnArgGluGluMetGluLeuGlnGly---GluGlnMetGlnArgGln 127
QY 380 AAGAGGAGCTCAGGAACCTTGCCTCAACAGTGGCGCTTAGGGCAACAGCGTTCGAG 439
Db MetArgGlyAlaGluAsnLeuLeuSerArgCysAsnLeu---SerProGlnArgCysPro 146
QY 440 TTG 442
Db 147 Met 147

RESULT 7
S01062
2S seed storage protein precursor (clone HaG5) - common sunflower

N/Alternate names: 2S albumin storage protein
C/Species: Helianthus annuus (common sunflower)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C/Accession: S01062
R/Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Tho
Mol. Gen. Genet. 210, 211-218, 1987
A/Title: Sequence and expression of a gene encoding an albumin storage protein in sunflo
A/Reference number: S01062; MUID:88142538; PMID:2830455
A/Accession: S01062
A/Molecule type: DNA
A/Residues: 1-295 <ALL>
A/Cross-references: UNIPROT:P15461; EMBL:X06410; NID:g18799; PIDN:CAA29699.1; PID:g18800
A/Note: part of this sequence, including the amino end of the mature protein, was confir
C/Genetics:
A/Intons: 192/2
C/Superfamily: gliadin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:162-295/Product: 2S seed storage protein #status experimental <MAT>

Alignment Scores:
Pred. No.: 3.93e-06 Length: 295
Score: 151.00 Matches: 47
Percent Similarity: 40.39% Conservative: 35
Best Local Similarity: 23.15% Mismatches: 65
Query Match: 11.94% Indels: 56
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x S01062 (1-295)
QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTCGC----- 55
Db 5 IleValLeuAlaLeuAlaPheAlaLeuValAlaPheAlaThrAlaHisThrThr 24
QY 56 -----AGCGACGAGTGGGAATCTCAAGGAGACAGAGATGCCAGAGCCAG 100
Db 25 IleThrThrThrIleGluAspGluAsnProIleSerGlyGlnArgGlnValSerGlnArg 44
QY 101 CTCGAGAGGCGCAACTGAGGCCCTGCGAGCAACATCTCTCATG-----CAGAAGATCCAA 154
Db 45 IleGlnGlyGlnArgLeuAsnGlnCysArgMetPheLeuGlnGlnGlyGlnAsnIlePro 64
QY 155 CGTGCAGGAGTTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCT 214
Db 65 ArgGluPheAspAsn----- 69
QY 215 AGTCCATATGATCGAGAGGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTGCAATGAG 274
Db 70 -----ProGlnMetGlyArgGlnGlnGlnGlnLeuGlnCysCysGlnGlu 86
QY 275 CTGAACGAGTTTGAGAACCAACCAAGGTGATGTCGAGGCAATGCAACAGATCATGAG 334
Db 87 LeuGlnAsnIleGluGly-----GlnCysGlnCysGluAlaValIysGlnValPheArg 104
QY 335 AACGAGCGATAGTTGTCAG-----GGGAGG 361
Db 105 GluAlaGlnGlnGlnValGlnGlnGlnGlnGlnGlnValProPheArgGlySer 124
QY 362 CAACGAGGACCAAGTTCAGAGGAGGAGCTTCAGGAACTTGCCTCAACAGTGGCGCTTAGG 421
Db 125 GlnGlnThrGlnGlnLeuIysGlnIysAlaGlnIleLeuProAsnValCysAsnLeuGln 144
QY 422 GCACACAGCGTTCGACCTTG-----GACGTCGAA 451
Db 145 Ser---ArgArgCysGluIleGlyThrIleThrThrThrValThrGluSerAsnIleAsp 163
QY 452 AGTGGCGGAGAGACAGATACCTAAACACCTCTCAAAAAAAGAAAAAGAAAGAA 511
Db 164 IleProPheArgAspArg-----ProPheGlyThrGlySerGlnGlnCysArgGlu 180
QY 512 AATAGCTTA 520
Db 181 ThrGluIle 183

RESULT 8
A29802
napin precursor (gNa) - rape
N/Alternate names: 1.7 S seed storage protein
C/Species: Brassica napus (rape)
C/Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 09-Jul-2004
C/Accession: A29802
R/Scofield, S.R.; Crouch, M.L.
J. Biol. Chem. 262, 12202-12208, 1987
A/Title: Nucleotide sequence of a member of the napin storage protein family from Brassi
A/Reference number: A29802; MUID:87308225; PMID:3040733
A/Accession: A29802
A/Molecule type: DNA
A/Residues: 1-186 <SCO>
A/Cross-references: UNIPROT:P09893; GB:J02782; NID:g167178; PIDN:AAA33007.1; PID:g167179
C/Genetics:
A/Gene: gNa
C/Superfamily: wheat alpha-amylase inhibitor

Alignment Scores:
Pred. No.: 9.49e-06 Length: 186
Score: 147.00 Matches: 47
Percent Similarity: 40.98% Conservative: 28
Best Local Similarity: 25.68% Mismatches: 52
Query Match: 11.62% Indels: 56
DB: 2 Gaps: 9

US-10-728-323-2 (1-717) x A29802 (1-186)

```

Qy 8 ATACTAGTACGCTCGCCCTTTCTCTCTCCGCCACGATCTGCG---AGGCAGCAG 64
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 ValSerAlaThrLeuAlaLeuPhePheLeuLeuThrAsnAlaSerValTyrArgThrVal 27
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 65 TGGGAATCCCAAGAGAC-----AGAGATGC 91
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 ValGluValAspGluAspAspAlaThrAsnProAlaGlyProPheArgLeuProLysCys 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 92 CAGAGCAGCTCGAGAGGCGC---AACCTGAGGCGCTCGAGCAACATCTCATCGACAAG 148
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGlnTrpLeuHisLysGln 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 149 ATCCAAGCTGAC-----160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 AlaMetGlnProGlyGlyGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPhe 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 161 GAGGATTCATATGAACGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 GluAspAspValGluAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 221 TATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAAC 280
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 -----GlnGlnCysCysAsnGluLeuHis 113
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 281 GAGTTTGAGAACCAAGGTGTCATGTGCGAGCATTG-----319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 GlnGluGlu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaVal 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 320 ---CAACAGATCATGGAGAACGAGCATAGTTGCGAGGAGGCAACAGGACACAG 376
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 ArgGlnValArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnVal 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 377 TTCAAGAGGAGGCTCAGC-----AACTTGCTCAACAGATGCGCGCTTAGGGACCA 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 IleSerArgValTyrGlnThrAlaThrHisLeuProArgValCysAsnIleArgGlnVal 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 428 CAGCGTTGC 436
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 SerIleCys 174
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9

```

NWML
2S albumin 1 precursor - Arabidopsis thaliana
N;Alternate names: seed storage protein At2S1
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: JAO161; PS0282; S34676; T06044
R;Krebbbers, E.; Herdies, L.; de Clercq, A.; Seurinck, J.; Leemans, J.; Van Damme, J.; Se
Plant Physiol. 87, 859-866, 1988
A;Title: Determination of the processing sites of an Arabidopsis 2S albumin and characte
A;Reference number: JAO161
A;Accession: JAO161
A;Molecule type: DNA
A;Residues: 1-164 <KRE>
A;Cross-references: UNIPROT:P15457; GB:M22032; NID:g166609; PIDN:AAA32743.1; PID:g166614
A;Accession: PS0282
A;Molecule type: protein
A;Residues: 38-73;84-162 <KR2>
R;Conceicao, A.D.S.; Krebbbers, E.
submitted to the EMBL Data Library, July 1993
A;Description: Tentative title: a cotyledon regulatory region is responsible for the dif
A;Reference number: S34674
A;Accession: S34676
A;Molecule type: DNA
A;Residues: 1-164 <CON>
A;Cross-references: EMBL:Z24745; NID:g395203; PIDN:CAA80870.1; PID:g395204
R;Bavan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Voickaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15484
A;Accession: T06044

```

```

A;Molecule type: DNA
A;Residues: 1-164 <BEV>
A;Cross-references: EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.90
A;Experimental source: cultivar Columbia; BAC clone T24A18
C;Genetics:
A;Gene: T24A18.90
A;Map position: 4
C;Superfamily: wheat alpha-amylase inhibitor
C;Keywords: seed; storage protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-164/Product: 2S albumin 1 proprotein #status predicted <AT2>
F;38-73/Product: 2S albumin 1 small chain #status experimental <SMC>
F;84-162/Product: 2S albumin 1 large chain #status experimental <LAC>

```

Alignment Scores:

Pred. No.:	1.06e-05	Length:	164
Score:	146.50	Matches:	46
Percent Similarity:	46.47%	Conservative:	33
Best Local Similarity:	27.06%	Mismatches:	52
Query Match:	11.58%	Indels:	39
DB:	1	Gaps:	9

US-10-728-323-2 (1-717) x NWML (1-164)

```

Qy 8 ATACTAGTACGCTCGCCCTTTCTCTCTCCGCCACGATCTGCG---AGGCAGCAG 64
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 ValCysAlaAlaLeuAlaLeuCysPheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 65 TGGGAATCCCAAGAGAC-----AGAAGATGCCAG 94
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 ValGluPheGluGluAspAlaThrAsnProIleGlyProLysMetArgLysCysArg 47
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 95 AGCCAGCTCGAGAGG---GCGAACCTGAGGCCCTCGAGCAACATCTCATGCAAGATC 151
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 LysGluPheGlnLysGluGlnHisLeuArgAlaCysGlnGlnLeuMetLeuGlnGlnAla 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 152 CAACGT-----CACGAGGATTCATATGAACGGAGCCCGTACAGCCCTAGTCAGGAT 202
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 ArgGlnLysArgSerGluPheAspGluPheGluAspMetGluAsnProGln----- 85
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 203 CCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGATCTCTCAGACACCAA----- 256
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 -----GlyGlnGlnGlnGlnGlnLeu 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 257 ---GAGAGGTGTTGCAATGAGTGAACGAGTTTGAACAACCAAGGTGTCATGTGCGAG 313
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 PheGlnGlnCysCysAsnGluLeuArgGlnGluGluProAsp-----CysValCysPro 111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 314 GCATTCCAAACAGATCATGGAGAACGAGCGGATAGTTGCGAGGAGGCAACAGAGCAAA 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 ThrLeuLysGlnAlaAla-----LysAlaValArgLeuGlnGlnHisGlnProMet 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 374 CAGTTCAAAGAGG-----GAGCTCAGGAACCTGCTCAACAGTGCAGCCTTAGGGCA 424
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 GlnValArgLysIleTyrGlnThrAlaLysHisLeuProAsnValCysAspIleProGln 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 425 CCACAGCGTTGCGACTTGGACGTCGAAAGT 454
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 ValAspValCysProPheAsnIleProSer 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10

```

S14946
2S seed storage protein large chain - Brazil nut
N;Alternate names: albumin 2S precursor
C;Species: Bertholletia excelsa (Brazil nut)
C;Date: 21-Nov-1993 #sequence revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S14946; S14479; S06252; S21640; B25802
R;Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Grossi
Plant Mol. Biol. 16, 437-448, 1991
A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin fr
A;Reference number: S14946; MUID:91370890; PMID:1840683
A;Accession: S14946
A;Status: preliminary

```



```

Db      37  HisGluAspGlyLeuTyrglyGluGluValGlnGlnArgSerCysGluGlnGln 56
      |||  |||:|||||  |||:|||||  |||:|||||  |||:|||||
QY     101  CTCGAGAGGGCGAACCTGGAGCCCTCGGAGCAACATCTCATGCAGAAAGATCCAACTGAC 160
      |||  |||  |||:|||||  |||:|||||  |||:|||||
Db      57  -----ArgLeuSerSerCysArgGluTyrrLeu-----GluArgPro 68
      |||  |||  |||:|||||  |||:|||||
QY     161  GAGGATTTCATATGAACGGGACCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
      |||  |||  |||:|||||
Db      69  Argasp----- 70
QY     221  TATGATCGGAGAGCGCTGGATCCTCTCAGACCAAGAGAGGTGTTCGAATGAGCTGAAC 280
      |||  |||  |||:|||||  |||:|||||  |||:|||||
Db      71  -----GlnProSerGluArgCysCysGluGluLeuGln 81
QY     281  GAGTTTGAGAACCAACAAAGGTGCATGTGGAGGCAATTCGAACAGATCATGAGAAC---- 337
      |||  |||  |||:|||||  |||:|||||  |||:|||||
Db      82  ArgMet-----SerProGlnCysArgCysGlnAlaIleGlnArgThrLeuGluAspVal 99
      |||  |||  |||:|||||  |||:|||||
QY     338  -----CAGAGCGAT-----AGG 349
      |||  |||  |||:|||||
Db      100  PheMetAspSerAspSerGlnAspGlyAlaProLeuAsnGlnArgArgGlnArgArg 119
      |||  |||  |||:|||||
QY     350  TTGCAGGGGAGG---CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTCCTCAA 406
      |||  |||  |||:|||||  |||:|||||
Db      120  GlyGlnGlyArgGlyMetGluGluGluGluValValArgAlaGluGluLeuProAsn 139
QY     407  CAGTGGCGCTTAGG---GCACACAGCGTTGGAGCTTGGAGCTCGAAAGT 454
      |||  |||  |||:|||||  |||:|||||
Db      140  ThrCysAsnValArgGlnSerProArgArgCysAspLeuGlnArgHisSer 156
      |||  |||  |||:|||||

RESULT 14
T08012
2S seed storage protein precursor - Douglas fir
C:Species: Pseudotsuga menziesii (Douglas fir)
C>Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T08012
R:Chatthai, M.; Mista, S.
Planta 206, 138-145, 1998
A:Title: Structure and expression of embryogenesis-specific cDNAs encoding 2S seed storage
A:Reference number: Z16286; MUID:98381325; PMID:9715536
A:Accession: T08012
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-170 <CHA>
A:Cross-references: UNIPROT:O64931; EMBL:AF029972; NID:G2988481; PIDN:AAC27000.1; PID:92
C:Genetics:
A:Gene: 2S3
C:Keywords: seed; storage protein

Alignment Scores:
Pred. No.: 0.00048 Length: 170
Score: 129.00 Matches: 43
Percent Similarity: 36.11% Conservative: 22
Best Local Similarity: 23.89% Mismatches: 49
Query Match: 10.20% Indels: 66
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x T08012 (1-170)
QY     8  ATACTAGTAGCCCTCGCCCTTTTCCTCGTCCGCCAGCATCTCGAGGCGAGCATGG 67
      |||  |||  |||:|||||  |||:|||||  |||:|||||
Db     17  ValSerLeuGlyValAlaLeuLeuLeuVal-----GlnTrp 29
QY     68  -----GAACTCCAAGGAGACAGAAGATGCCAGAGC 97
      |||  |||  |||:|||||
Db     30  SerThrProAsnValAspAlaAlaGlyAspAsnMetPheGlyGluAspValValGlnGln 49
QY     98  CAGCTCGAGAGGGCGAACCTTGAGGCCCTCGGACCAACATCTCATGCAGAAAGATCCAACGT 157
      |||  |||  |||:|||||
Db     50  GlnGlnArgArgGlySerCysAspPro-----GlnArgLeuSerSer 63
QY     158  GACGAGGATTTCATATGAACGGGACCGCTAGTACGCTTAGTCAGGATCCGTACAGCCCTAGT 217
```

```

Db      64  CysArgAspTyrrLeuGluArgArgGluGlnProSer----- 76
      |||  |||  |||:|||||  |||:|||||
QY     218  CCATATGATCGGAGAGGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTCGAATGAGCTG 277
      |||  |||  |||:|||||  |||:|||||
Db      77  -----GluArgCysCysAsnGluLeu 83
QY     278  AACGAGTTTGAGAACCAACCAAGGTGCATGTGCGAGGCAATTCGAACAGATCATGAGAAC 337
      |||  |||  |||:|||||  |||:|||||
Db      84  GluArgMet-----SerProGlnCysArgCysProAlaIleGlnGlnValPheAspGln 101
      |||  |||  |||:|||||
QY     338  CAGAGCGATGAGTTG----- 352
Db     102  SerSerGluAspLeuSerMetValAspSerHisSerGlnAsnAlaAlaGlyAsnGlnArg 121
QY     353  -----CAGGGAGGCAACAG---GAGCAACAGTTCAAGAGGAGCTCAGGAAC 397
      |||  |||  |||:|||||  |||:|||||
Db     122  ArgArgGluGluArgGlyArgGluGluAlaGluGluMetValGluArgAlaGlnArg 141
QY     398  TTGCCTCAACAGTGCAGCGCTTAGGGCACACAGCGT---TGGAGCTTGGAGCTCGAAAGT 454
      |||  |||  |||:|||||  |||:|||||
Db     142  LeuProAsnThrCysAsnValArgGlnProProArgHisCysAspIleGlnArgHisSer 161

RESULT 15
S18871
2S-like storage protein - white spruce
C:Species: Picea glauca (white spruce)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S18871
R:Newton, C.H.
submitted to the EMBL Data Library, November 1991
A:Reference number: S18871
A:Accession: S18871
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-172 <NEW>
A:Cross-references: UNIPROT:P26986; EMBL:X63193; NID:G20495; PID:G20496

Alignment Scores:
Pred. No.: 0.000596 Length: 172
Score: 128.00 Matches: 46
Percent Similarity: 37.43% Conservative: 21
Best Local Similarity: 25.70% Mismatches: 46
Query Match: 10.12% Indels: 66
DB: 2 Gaps: 9

US-10-728-323-2 (1-717) x S18871 (1-172)
QY     14  GTAGCCCTCGCCCTTTTCCTCCTC-----GCTGCCCAACGCA 49
      |||  |||  |||:|||||  |||:|||||
Db     18  LeuSerValAlaLeuPheLeuLeuPheHisTrpGlyIleProSerValAspGlyHisGlu 37
QY     50  TCTGCGAGCGACAGTGGGAATCCAAAGGAGACAGAAGA---TGCAGAGCCAGCTCGAG 106
      |||  |||  |||:|||||  |||:|||||
Db     38  AspAsnMetTyrrGlyGluGluIleGlnGlnArgSerCysAspProGlnArgasp 57
QY     107  AGGGCGAACCTTGAGCGCCCTCGAGGACACATCTCATGCAGAAAGATCCAACGTGACGAGAT 166
      |||  |||  |||:|||||
Db     58  ProGlnArgLeuSerSerCysArgAspTyrrLeu----- 68
QY     167  TCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGAT 226
      |||  |||  |||:|||||
Db     69  -----GluArgArgGluGlnProSer----- 76
QY     227  CGGAGAGGCGCTGGATCCTCTCAGCAACCAAGAGAGGTGTTCGAATGAGTGAACAGTGT 286
      |||  |||  |||:|||||  |||:|||||
Db     77  -----GluArgCysCysGluGluLeuGlnArgMet 86
QY     287  GAGAACCAACAAAGGTGCATGTGCGAGGCAATTCGAACAGATCATGAG----- 334
      |||  |||  |||:|||||  |||:|||||
Db     87  -----SerProGlnCysArgCysGlnAlaIleGlnGlnMetLeuSerSerLeuSer 104
QY     335  -----AACAGAGCGATAGG 349
```

```

Db      105 TyrAspSerPheMetAspSerAspSerGlnGluAspAlaProLeuAsnGlnArgArgArg 124
Qy      350 TTG-----CAGGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACTTG 400
Db      125 ArgArgGluGlyArgGlyArgGluGluGluAlaMetGluArg---AlaAlaTyrLeu 143
Qy      401 CCTCAACAGTCGGCCTTAGG---GCACCACAGCGTTGCGACTTGGACGTCGAAAGT 454
Db      144 ProAsnThrCysAsnValArgGluProProArgArgCysAspIleGlnArgHisSer 162

```

Search completed: August 24, 2005, 10:06:37
Job time : 11.9772 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 43.8791 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcacatactagtagcccc.....taaaagaatcatgttttgggt 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10728323/runat_23082005_124355_29210/app_query.fasta.1.4757
-DB=uniprot 03 -QFMT=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10728323 @CGN 1.1 291 @runat_23082005_124355_29210 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	842	66.6	179	2	Q647H0 arachis hyp
2	838	66.2	175	2	Q6PSU1 arachis hyp
3	819	64.7	156	2	Q941R0 arachis hyp
4	815	64.4	172	2	Q8GV20 arachis hyp
5	815	64.4	187	2	Q6PSU2 arachis hyp
6	808	63.9	169	2	Q7YLC0 arachis hyp
7	792	62.6	166	2	Q84TU1 arachis hyp
8	454.5	35.9	145	2	Q647G9 arachis hyp
9	443.5	35.1	144	2	Q8W251 arachis hyp
10	388.5	30.7	129	2	Q9SQG5 arachis hyp
11	300	23.7	160	2	Q9SQH1 arachis hyp
12	299	23.6	158	2	Q647G8 arachis hyp
13	278	22.0	153	2	Q99235 lupinus ang
14	243	19.2	158	1	P19594 glycine max
15	241.5	19.1	155	2	Q9ZN74 glycine max
16	187.5	14.6	80	1	CG2L_LUPAN lupinus ang

17	184.5	14.6	323	2	Q39928 helianthus
18	177.5	14.0	168	2	Q8LPD3 linum usita
19	173.5	13.7	141	2	Q39649 cucurbita c
20	167	13.2	160	2	Q71HN1 ficus awkeo
21	164.5	13.0	148	2	Q9XHP1 sesamum ind
22	159.5	12.6	143	2	Q8VX62 sesamum ind
23	159	12.6	161	2	Q7YLC2 juglans nig
24	153.5	12.1	139	2	P93198 juglans reg
25	153	12.1	153	2	Q9AUD1 sesamum ind
26	151	11.9	295	1	2SS5 HELAN
27	149.5	11.7	167	2	Q84JW2 vitis vinif
28	148.5	11.7	140	2	Q8L694 momordica c
29	148	11.7	285	2	Q8GUD8 helianthus
30	147	11.6	186	1	2SS5 BRANA
31	146.5	11.6	164	1	2SS1 ARATH
32	144.5	11.4	165	2	Q9FHP31 arabidopsis
33	144	11.4	169	2	Q8LPD4 linum usita
34	143.5	11.3	179	2	Q84NG9 vitis vinif
35	141.5	11.2	138	2	Q8H2B8 anacardium
36	141	11.1	126	2	Q8RW54 helianthus
37	140.5	11.1	146	1	2SS BEREX
38	139	11.0	155	1	2SS2 CAPMA
39	139	11.0	184	2	Q42444 brassica ma
40	139	11.0	258	1	2SS RICCO
41	132.5	10.5	165	2	Q40850 picea glauc
42	132.5	10.5	167	2	Q81411 picea glauc
43	130.5	10.3	161	2	Q81410 picea glauc
44	129.5	10.2	173	2	Q81412 picea glauc
45	129	10.2	170	2	Q64931 pseudotsuga

ALIGNMENTS

RESULT 1
Q647H0 PRELIMINARY; PRT; 179 AA.
AC Q647H0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE 2S protein 1 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV722689; AAU21494.1; -;
FT NON TER 1
SQ SEQUENCE 179 AA; 20850 MW; 58CE6453B409E1DC CRC64;

Alignment Scores:
Pred. No.: 2,57e-65 Length: 179
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q647H0 (1-179)

QY	2	CTCACATACTAGTAGCCCTTCCTCTCGCTGCCACGCATCTGGAGGACG	61
DB	23	LeuThrIleuValAlaLeuAlaLeuLeuLeuAlaHisAlaSerAlaArgGln	42
QY	62	CAGTGGGAACCTCCAAAGGAGACAGACAGCTCGAGAGGGGGAACCTGAGG	121
DB	43	GlntrpGluLeuGlnGlyAspArgGlyGlnSerGlnLeuGluArgAlaAsnLeuArg	62

QY	122	CCCTGCGAGCAACATCTCATGCGAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	63	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	82
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241
Db	83	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly	102
QY	242	TCCTCTCAGACCAAGAGAGGTGTCATGACCTGAACAGATTGAGAACACCAAGG	301
Db	103	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	122
QY	302	TGCATGTGCGAGGATTGCAACAGATCATCGAGAACCCAGAGCGATAGTTGCGAGGGAGG	361
Db	123	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	142
QY	362	CAACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTCGCTCAACAGATGCGCGCTTAGG	421
Db	143	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	162
QY	422	GCACACAGCGTTGCGACTTGGAGCTGCAAGTGGCGGCGGAGACAGATAC	472
Db	163	AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr	179
RESULT 2			
ID	Q6PSU1	PRELIMINARY;	PRT; 175 AA.
AC	Q6PSU1; 2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)		
DE	Seed storage protein SSP2 (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]_TaxID=3818;		
RP	SEQUENCE FROM N.A.		
RA	Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;		
RL	Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY581854; AAT0599.1; -.		
DR	InterPro; IPR003612; AAI.		
DR	Pfam; PF00234; Tryp_alpha_amyl; 1.		
DR	SMART; SM00499; AAI; 1.		
FT	NON_TER 1 1		
SQ	SEQUENCE 175 AA; 20463 MW; B3FB5806B70733DE CRC64;		
Alignment Scores:			
Pred. No.:	5.77e-65	Length:	175
Score:	836.00	Matches:	156
Percent Similarity:	99.36%	Conservative:	0
Best Local Similarity:	99.36%	Mismatches:	1
Query Match:	66.25%	Indels:	0
DB:	2	Gaps:	0
US-10-728-323-2 (1-717) x Q6PSU1 (1-175)			
QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTCCGCCACGCATCTGCGAGGAG	61
Db	19	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln	38
QY	62	CAGTGGGAACCTCAAGAGACAGAGATGCCAGGCGAGCTCGAGAGGGCGAACCCTGAG	121
Db	39	GlnTrpGluPheGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	58
QY	122	CCCTGCGAGCAACATCTCATGCGAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	59	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	78
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241

Db	79	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly	98
QY	242	TCCTCTCAGACCAAGAGAGGTGTCATGACCTGAACAGATTGAGAACACCAAGG	301
Db	99	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	118
QY	302	TGCATGTGCGAGGATTGCAACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAGG	361
Db	119	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	138
QY	362	CAACAGAGCAACAGTTCAAGAGGGAGCTCAGAACTTCGCTCAACAGATGCGCGCTTAGG	421
Db	139	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	158
QY	422	GCACACAGCGTTGCGACTTGGAGCTGCAAGTGGCGGCGGAGACAGATAC	472
Db	159	AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr	175
RESULT 3			
ID	Q941R0	PRELIMINARY;	PRT; 156 AA.
AC	Q941R0; 2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Allergen II (Fragment).		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]_TaxID=3818;		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21192158; PubMed=11295663; DOI=10.1067/mai.2001.113522;		
RA	Viquez O.M., Summer C.G., Dodo H.W.;		
RT	"Isolation and molecular characterization of the first genomic clone		
RT	of a major peanut allergen, Ara h 2.;"		
RL	J. Allergy Clin. Immunol. 107:713-717(2001).		
DR	EMBL; AY007229; AAK96887.1; -.		
DR	InterPro; IPR003612; AAI.		
DR	Pfam; PF00234; Tryp_alpha_amyl; 1.		
DR	SMART; SM00499; AAI; 1.		
FT	NON_TER 156 156		
SQ	SEQUENCE 156 AA; 18167 MW; 664D6832F91F36D1 CRC64;		
Alignment Scores:			
Pred. No.:	2.68e-63	Length:	156
Score:	819.00	Matches:	153
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	64.74%	Indels:	0
DB:	2	Gaps:	0
US-10-728-323-2 (1-717) x Q941R0 (1-156)			
QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTCCGCCACGCATCTGCGAGGAG	61
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln	23
QY	62	CAGTGGGAACCTCAAGAGACAGAGATGCCAGGCGAGCTCGAGAGGGCGAACCCTGAG	121
Db	24	GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg	43
QY	122	CCCTGCGAGCAACATCTCATGCGAGAGATCCAAAGTGCAGGATTCATATGAACGGAC	181
Db	44	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	63
QY	182	CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGAGAGCGCTGGA	241
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly	83

```

QY 242 TCCTCTCAGCACCAAGAGGCTTGCATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluAArgCysCysAsnGluLeuAsnGluPheGluAAsnGlnArg 103
QY 302 TGCATGTGCGAGGCAATTCGAACAGATCATGGAGAACAGAGCGATAGTTTGCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCTCAACAGTCGGCGCTTAGG 421
Db 124 GlnGlnGlnGlnGlnPheGlyAspGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAAGAGGCTTGGAGCTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 4
ID Q8GV20 PRELIMINARY; PRT; 172 AA.
AC Q8GV20;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Allergen Ara h 2.02.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Chatel J.-M., Bernard H., Orson F.M.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY158467; AAN7576.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 172 AA; 20114 MW; 8BB991C8D8C143AB CRC64;

Alignment Scores:
Pred. No.: 6.04e-63 Length: 172
Score: 815.00 Matches: 155
Percent Similarity: 92.31% Conservative: 1
Best Local Similarity: 91.72% Mismatches: 1
Query Match: 64.43% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q8GV20 (1-172)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCCCAACGATCTGCAGGCAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGNACTCCAAGGAGACAGAGATGCCAGAGCAGCTCCAGAGGGCGAACCTGAG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGCAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGlyArgAsp 63
QY 182 CCTACAGCCCTAGTCAG-----GATCCG 205
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 83
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGTGT 265
Db 84 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 103
QY 266 TGCATAGTGAACGAGTTTGAGAACCAACAAAGGTGATGTCGAGGCGATTGCAACAG 325
Db 104 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 123

```

```

QY 326 ATCATCGAGAACCAAGAGCGATAGTTGCAGGGAGGCAACAGAGCAACAGTTCAAGAGG 385
Db 124 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlyArgGlnGlnGlnPheGlyArg 143
QY 386 GAGCTCAGAACTTGCTCAACAGTCGGCTTTCAGGGCCACCAAGCGTTTGCAGCTTGGAC 445
Db 144 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 163
QY 446 GTCGAAAGTGGCGGCAGAGACATAC 472
Db 164 ValGluSerGlyGlyArgAspArgTyr 172

RESULT 5
ID Q6PSU2 PRELIMINARY; PRT; 187 AA.
AC Q6PSU2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Seed storage protein SSPI (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581853; AAT00598.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
DR NON_TER 1.
SQ SEQUENCE 187 AA; 21786 MW; 850AF22C94983FB8 CRC64;

Alignment Scores:
Pred. No.: 6.07e-63 Length: 187
Score: 815.00 Matches: 155
Percent Similarity: 92.31% Conservative: 1
Best Local Similarity: 91.72% Mismatches: 1
Query Match: 64.43% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q6PSU2 (1-187)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCCCAACGATCTGCAGGCAG 61
Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 38
QY 62 CAGTGGGNACTCCAAGGAGACAGAGATGCCAGAGCAGCTCCAGAGGGCGAACCTGAGG 121
Db 39 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58
QY 122 CCTCGCAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTCATATGAACGGAC 181
Db 59 ProCysGluGlnHisLeuMetGlnIleGlnArgAspGluAspSerTyrGlyArgAsp 78
QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
Db 79 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 98
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGTGT 265
Db 99 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 118
QY 266 TGCATAGTGAACGAGTTTGAGAACCAACAAAGGTGATGTCGAGGCGATTGCAACAG 325
Db 119 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 138
QY 326 ATCATCGAGAACCAAGAGCGATAGTTGCAGGGAGGCAACAGAGCAACAGTTCAAGAGG 385

```

```

|||||
Db 139 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlnGlnGlnPheLysArg 158
|||||
QY 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGGTTGCCAGCTTGGAC 445
|||||
Db 159 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 178
|||||
QY 446 GTCGAAAGTGGCGGCAGACAGATAC 472
|||||
Db 179 ValGluSerGlyGlyArgAspArgTyr 187
|||||

RESULT 6
Q7Y1C0 PRELIMINARY; PRT; 169 AA.
AC Q7Y1C0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 26, Last annotation update)
DE Allergen Ara h 2 isoform (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX Paik-Ro O.G., Seib J.C., Smith R.L.;
RT "Seed-specific, developmentally regulated genes of peanut.";
RL Theor. Appl. Genet. 104:236-240(2002).
DR EMBL; AY117434; AAM78596.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; TYP_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 169 AA; 19768 MW; 0B14A7ED911F34EF CRC64;

```

```

Alignment Scores:
Pred. No.: 2,49e-62 Length: 169
Score: 808.00 Matches: 154
Percent Similarity: 91.72% Conservative: 1
Best Local Similarity: 91.12% Mismatches: 2
Query Match: 63.87% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q7Y1C0 (1-169)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACCGCATCTCGAGGAG 61
|||||
Db 1 LeuThrIleLeuValAlaProAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
|||||
QY 62 CAGTGGGAACCTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
|||||
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
|||||
QY 122 CCTCGCAGCAACATCTCATGCAAGAATCCAACTGACGAGGATTCATATGAACGGGAC 181
|||||
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 60
|||||
QY 182 CGGTACAGCCCTAGTCAG-----GATCCG 205
|||||
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgArgAspPro 80
|||||
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
|||||
Db 81 TyrSerProSerProTyrAspArgArgGlyAlaGlySerSerGlnHisGlnGluArgCys 100
|||||
QY 266 TGCATAGCTGAACGAGTTTGAGAACCAACAAAGGTGCTATGTCGAGGCATTCGAACAG 325
|||||
Db 101 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 120
|||||
QY 326 ATCATCGAGACACAGCGGATAGTTGTCAGGGGAGGCAACAGGCAACAGTTCAAGAGG 385
|||||
Db 121 IleMetGluAsnGlnSerAspArgLeuGlnGlnGlnGlnGlnPheLysArg 140
|||||

```

```

QY 386 GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCCACACAGGTTGCCAGCTTGGAC 445
|||||
Db 141 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 160
|||||
QY 446 GTCGAAAGTGGCGGCAGACAGATAC 472
|||||
Db 161 ValGluSerGlyGlyArgAspArgTyr 169
|||||

RESULT 7
Q84TU1 PRELIMINARY; PRT; 166 AA.
AC Q84TU1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Seed storage protein (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX Paik-Ro O.G., Seib J.C., Smith R.L.;
RT "Seed-specific, developmentally regulated genes of peanut.";
RL Theor. Appl. Genet. 104:236-240(2002).
DR EMBL; AF366560; AAO61750.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; TYP_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 166 AA; 19321 MW; BD04F7F26CE7B437 CRC64;

Alignment Scores:
Pred. No.: 6,31e-61 Length: 166
Score: 792.00 Matches: 151
Percent Similarity: 92.12% Conservative: 1
Best Local Similarity: 91.52% Mismatches: 1
Query Match: 62.61% Indels: 12
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x Q84TU1 (1-166)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACCGCATCTCGAGGAG 61
|||||
Db 2 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 21
|||||
QY 62 CAGTGGGAACCTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
|||||
Db 22 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 41
|||||
QY 122 CCTCGCAGCAACATCTCATGCAAGAATCCAACTGACGAGGATTCATATGAACGGGAC 181
|||||
Db 42 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGlyArgAsp 61
|||||
QY 182 CGGTACAGCCCTAGTCAG-----GATCCG 205
|||||
Db 62 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgArgAspPro 81
|||||
QY 206 TACAGCCCTAGTCATATGATCGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
|||||
Db 82 TyrSerProSerProTyrAspArgArgGlyAlaGlySerSerGlnHisGlnGluArgCys 101
|||||
QY 266 TGCATAGCTGAACGAGTTTGAGAACCAACAAAGGTGCTATGTCGAGGCATTCGAACAG 325
|||||
Db 102 CysAsnGluLeuAsnGluPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 121
|||||
QY 326 ATCATGGAGAACACAGCGGATAGTTGTCAGGGGAGGCAACAGGCAACAGTTCAAGAGG 385
|||||

```



```

AC Q647G8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE 2S protein 2.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Aeschyls.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722691; AAU21496.1;
SQ SEQUENCE 158 AA; 18455 MW; 32AD4D4061544948 CRC64;

Alignment Scores:
Pred. No.: 1.25e-17 Length: 158
Score: 299.00 Matches: 75
Percent Similarity: 63.12% Conservatives: 26
Best Local Similarity: 46.88% Mismatches: 39
Query Match: 23.64% Indels: 21
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q647G8 (1-158)
QY 2 CTCACCATAGTAGAGCCCTC-----GCCCTTTTCTCCTCGCTGCC----- 43
DB 4 LeuSerIleuValAlaLeuLeuGlyAlaLeuValValAlaSerAlaThrArgTyr 23
QY 44 -----CACGCATCTCGAGCAGCAGTGGAACTC-----CAAGGAGCAGAAGA 88
DB 24 AspProAspArgGlySerArgGlyLeuArgTyrAlaProSerArgGlyAspGln 43
QY 89 TGCAGAGCCAGCTCGAGAGGCGCAACTGAGGCCCTCGAGCAACATCTCATGCGAAG 148
DB 44 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHisIleArgGlnArg 63
QY 149 ATCCAACTGACGAGGATTCATATGAAACGGGACCGCTACAGCCCTAGTCAGGATCCGTAC 208
DB 64 ValGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
QY 209 AGCCCT-----AGTCCATATGATCGGAGAGGCGCTCGATCTCTCTACACCAAGAGAGG 262
DB 78 SerGlnArgGlySerArgGlyArgProGlyGluSerAspGluAspGlnGlnArg 97
QY 263 TGTGCAATGAGCTGACGAGTGTGAGAACACCAAGGTGTCATGTCGAGGCGATTGCAA 322
DB 98 CysCysAsnGluLeuAsnArgPheGlnAsnAsnGlnArgCysMetCysGlnAlaLeuGln 117
QY 323 CAGATCATGAGAACACGAGATAGTGTG---CAGGGAGGCAACAGGAGCAACAGTT- 378
DB 118 GlnIleuGlnAsnGlnSerPheArgPheGlnGlnAspArgSerGlnLeuHisGlnAsn 137
QY 379 CAAGAGGAGGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCACACAGCGTTCGGA 438
DB 138 GlyGluGlyAlaGlnGluLeuAlaProGluLeuArgValGlnValThrLysProLeuArg 157

RESULT 13
Q99235 PRELIMINARY; PRT; 153 AA.
AC Q99235;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Conglutin delta.
OS Lupinus angustifolius (Narrow-leaved blue lupine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.

OX NCBI_TaxID=3871;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Unicrop; TISSUE=Cotyledon;
RX MEDLINE=91355912; PubMed=2103479;
RA Gayler K.R., Kollivas S., Macfarlane A.J., Lilley G.G., Baldi M.,
RA Blagrove R.J., Johnson E.D.;
RT "Biosynthesis, cDNA and amino acid sequences of a precursor of
RT conglutin delta, a sulphur-rich protein from Lupinus angustifolius.";
RL Plant Mol. Biol. 15:879-893(1990).
DR EMBL; X53523; CAA37598.1;
DR PIR; S12404; A33090.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 153 AA; 17785 MW; A45B7025353F0769 CRC64;

Alignment Scores:
Pred. No.: 8.7e-16 Length: 153
Score: 278.00 Matches: 62
Percent Similarity: 55.00% Conservatives: 26
Best Local Similarity: 38.75% Mismatches: 46
Query Match: 21.98% Indels: 26
DB: 2 Gaps: 5

US-10-728-323-2 (1-717) x Q99235 (1-153)
QY 2 CTCACCATAGTAGAGCCCTC---GCCCTTTTCTCCTCGCTGCCACGCAATCTGCAGG 58
DB 4 LeuThrIleuIleAlaLeuValAlaLeuValValValHisThrSerAlaPhe 23
QY 59 CACAGTGGGAATCTCCAGGAGACAGAGATGTCAGAGCCAGCTCGAGAGGGCGAACCTG 118
DB 24 Gln-----SerSerLysGlnSerCysLysArgGlnGlnGlnValAsnLeu 39
QY 119 AGCCCTTGCAGCAACATCTCATGAGAGATCAACGTCAGGAGGATTCATATGAACGG 178
DB 40 ArgHisCysGluAsnHisIleAlaGlnArgIleGlnGlnGlnGlnGlnGluAsp 59
QY 179 GACCCGTACAGC-----CTATGTCAGGATCCG 205
DB 60 HisAlaLeuLysLeuArgGlyIleLysHisValIleLeuArgHisArgSerGlnGlu 79
QY 206 TACAGCCCTAGTCCATATGATCGAGAGGCGCTGATCTCTCAGCACCAAGAGAGTGT 265
DB 80 TyrSerGlu-----GluSerGluGluLeuAspGlnCys 90
QY 266 TGCAAATGAGCTGAACGAGTTTGAGAACACAAAGGTGCATGTGCGAGGCAATTCACACAG 325
DB 91 CysGluGlnLeuAsnGlnLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGln 109
QY 326 ATCATGGAGAACACGAGCGATAGTTTGCAGGGAGGCAACAGAGCAACAGTTTCAAGAGG 385
DB 110 IleTyrGluSerGlnSerGluGlnCysGluGlySerGlnGlnGlnGlnGlnGlnGln 129
QY 386 GACCTCAGGAACTTGCCTCAACAGTGGCGCTTAGGGCACACAGCGTTCGAGATTGGAC 445
DB 130 GluLeuGluLysLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValAsn 149

RESULT 14
2SS_SOYBN STANDARD; PRT; 158 AA.
AC P19594;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 2S albumin precursor (G2S-1) [Contains: 2S albumin small chain
DE (Aspartic acid-rich peptide); 2S albumin large chain (8 kDa
DE methionine-rich protein) (8 kDa MRP)].
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

```

```

OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv, Hodgson 78; TISSUE=Cotyledon;
RA Galvez A.F., Revilla M.J.R., de Lumen B.O.;
RT "A novel methionine-rich protein from soybean cotyledon: cloning and
characterization of cDNA.";
RL (er) Plant Gene Register PGR97-103.
RN [2]
RP SEQUENCE OF 22-64.
RC TISSUE=Seed;
RA MEDLINE=87280104; PubMed=3611081;
RA Odani S., Koide T., Ono T.;
RT "Amino acid sequence of a soybean (Glycine max) seed polypeptide
having a poly(L-aspartic acid) structure.";
RL J. Biol. Chem. 262:10502-10505 (1987).
RN [3]
RP SEQUENCE OF 82-96.
RA Revilla M.J., Galvez A.F., Krenz D.C., de Lumen B.O.;
RT "An 8 kDa methionine-rich protein (MRP) from soybean (Glycine max)
cotyledon: identification, purification and N-terminal sequence.";
RL J. Agric. Food Chem. 44:2930-2935 (1996).
CC -1- FUNCTION: This is a 2S seed storage protein.
CC -1- SUBUNIT: The protein consists of two chains linked by 2 disulfide
bonds.
CC -1- SIMILARITY: Belongs to the 2S seed storage albumins family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF005030; AAB71140.1; -.
DR PIR; T05710; T05710.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
KW Direct protein sequencing; Seed storage protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 64 2S albumin small chain.
FT PROPEP 65 81
FT CHAIN 82 158 2S albumin large chain.
FT SITE 54 56 Cell attachment site (Potential).
FT DOMAIN 56 64 Poly-Asp.
FT DOMAIN 88 91 Poly-Glu.
SQ SEQUENCE 158 AA; 18460 MW; C1F42723B0F62D81 CRC64;

Alignment Scores:
Pred. No.: 1,03e-12 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.49% Conservative: 31
Best Local Similarity: 36.36% Mismatches: 55
Query Match: 19.21% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x 2SS_SOYBN (1-158)

QY 5 ACCATACTAGTACCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGCGACG 64
Db 5 ThrileuLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGGCC 124
Db 24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
QY 125 TGGGAGCAACATCTCATGCAAGAGATCCAA-----CGTCAGCAGGATTCATATGAACGG 178
Db 43 CysGlnHisIleMetGlnLysIleGlnGlyArgGlyAspAspAspAsp 62
QY 179 GACCCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGGCT 238
```

```

Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCTCTCTCAG-----CACCAGAGAGGCTGTTCATGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGGAGAACCAACCAAGGTGCATGTCCGAGGATTCGAACAGATCATGAGAACACG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGCTTCGAGGGAGGCAAGGAGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGlnLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGGCGCTTAGGGCCACACAGAGCTTCGGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 15
Q9ZNZ4
ID Q9ZNZ4 PRELIMINARY; PRT; 155 AA.
AC Q9ZNZ4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Napin-type 2S albumin 1 precursor.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]_TaxID=3847;
RP SEQUENCE FROM N.A.
RC STRAIN=variety PIONEER 9341; TISSUE=Seed;
RA Jung R., Hastings C., Coughlan S.J., Hu W.-N.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U71194; AAD09630.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
KW Signal.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 155 napin-type 2S albumin 1.
SQ SEQUENCE 155 AA; 17835 MW; 2C15E5E936301E3 CRC64;

Alignment Scores:
Pred. No.: 1.4e-12 Length: 155
Score: 241.50 Matches: 58
Percent Similarity: 57.14% Conservative: 30
Best Local Similarity: 37.66% Mismatches: 53
Query Match: 19.09% Indels: 13
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q9ZNZ4 (1-155)

QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCATCTCGAGGCGAC 61
Db 4 LeuThrIleLeuLeuIleAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer 21
QY 62 CAGTGGGAACCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCGGACCTGAGG 121
Db 22 LysTrpGlnGlnHisGlnGlnGluSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41
QY 122 CCCTGGGACCAATCTCATGCAAGAGATCCAA-----CGTCAGCAGGATTCATAT 172
Db 42 ProCysGlu---HisIleMetGlnLysIleGlnAlaGlyArgArgGlyGluAspGlySer 60
QY 173 GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGA 232
Db 61 AspGluAspHisIleLeuIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 79
QY 233 GCGCTGGATCTCTCAG-----CACCAGAGAGGCTGTTCATGAGCTGAAC 280
```

```

Db      80  LysGluGlyLysGluGluGluGluGluGlyHisMetGlnLysCysSerGluMetSer 99
Qy      281  GAGTTTGAGAACAAACCAAGGTGCATGTGCGAGCATTTGCAACAGATCATGAGAACCCAG 340
Db      100  GluLeuLys--SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln 118
Qy      341  AGCGATAGGTTGCAGGGGAGGCAACAGGAGCAGTTCAGAGGGAGCTCAGGAACCTTG 400
Db      119  SerGluGlnLeuGluGlyLys--GluLysLysGlnMetGluArgGluLeuMetAsnLeu 137
Qy      401  CCTCAACAGTGGCGCCTTAGGGCACACAGCGTTCGACTTG 442
Db      138  AlaIleArgCysArgLeuGlyProMetIleGlyCysAspLeu 151

```

Search completed: August 24, 2005, 10:04:45
 Job time : 47.8791 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 09:03:06 ; Search time 12.417 Seconds
(without alignments)
8620.958 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcacactactagtagcc.....taaaagatcatgttttgg 717

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124356_29238/app_query.fasta_1.4757
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1_46 @runat_23082005_124356_29238 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	66.6	157	4	US-09-106-872A-22
2	842	66.6	157	4	US-09-191-593-10
3	842	66.6	157	4	US-09-106-872A-2
4	827	65.4	157	4	US-09-191-593-21
5	243	19.2	158	2	US-08-618-911-4
6	243	19.2	158	3	US-08-938-675A-2
7	243	19.2	158	4	US-09-531-727-2
8	241.5	19.1	155	2	US-08-618-911-2
9	229	18.1	158	2	US-08-618-911-6
10	156	12.3	28	4	US-09-191-593-53
11	156	12.3	158	2	US-08-670-186-4
12	151	11.9	158	2	US-08-670-186-6
					Sequence 22, Appli
					Sequence 10, Appli
					Sequence 2, Appli
					Sequence 21, Appli
					Sequence 4, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 6, Appli
					Sequence 53, Appli
					Sequence 4, Appli
					Sequence 6, Appli

13	139	11.0	155	2	US-08-670-186-2	Sequence 2, Appli
14	139	11.0	184	1	US-08-453-924-3	Sequence 3, Appli
15	138	10.9	28	4	US-09-191-593-58	Sequence 58, Appli
16	136.5	10.8	164	4	US-09-627-536-13	Sequence 13, Appli
17	118	9.3	30	4	US-09-106-872A-16	Sequence 16, Appli
18	118	9.3	30	4	US-09-191-593-57	Sequence 57, Appli
19	106.5	8.4	174	4	US-09-645-593-7	Sequence 7, Appli
20	106.5	8.4	1088	4	US-09-233-857-13	Sequence 13, Appli
21	105.5	8.3	1099	3	US-09-442-100-2	Sequence 2, Appli
22	105.5	8.3	1099	4	US-08-939-106-2	Sequence 2, Appli
23	105.5	8.3	1099	4	US-09-442-102-2	Sequence 2, Appli
24	102	8.1	316	3	US-08-098-327B-31	Sequence 31, Appli
25	102	8.1	316	3	US-08-462-625-31	Sequence 31, Appli
26	100	7.9	17	4	US-09-191-593-55	Sequence 55, Appli
27	100	7.9	17	4	US-09-191-593-61	Sequence 61, Appli
28	100	7.9	17	4	US-09-191-593-62	Sequence 62, Appli
29	100	7.8	224	4	US-09-270-767-44901	Sequence 44901, A
30	99.5	7.9	148	4	US-09-083-852A-4	Sequence 4, Appli
31	99.5	7.9	148	4	US-09-083-852A-6	Sequence 6, Appli
32	99.5	7.9	148	4	US-09-489-674B-6	Sequence 6, Appli
33	97	7.7	185	4	US-09-252-991A-23490	Sequence 23490, A
34	95	7.5	32	1	US-08-158-704-1	Sequence 1, Appli
35	95	7.5	32	2	US-08-610-424B-1	Sequence 1, Appli
36	95	7.5	32	4	US-09-336-463-1	Sequence 1, Appli
37	95	7.5	32	4	US-09-191-593-1	Sequence 1, Appli
38	95	7.5	152	4	US-09-252-991A-19284	Sequence 19284, A
39	94	7.4	43	3	US-09-303-814-1	Sequence 1, Appli
40	94	7.4	500	4	US-09-265-630-13	Sequence 13, Appli
41	93.5	7.4	415	4	US-09-252-991A-18252	Sequence 18252, A
42	93	7.4	537	4	US-09-252-991A-27024	Sequence 27024, A
43	92.5	7.3	788	2	US-08-918-914-4	Sequence 4, Appli
44	91.5	7.2	1898	1	US-08-056-200-94	Sequence 94, Appli
45	91.5	7.2	1898	2	US-08-800-644-94	Sequence 94, Appli

ALIGNMENTS

RESULT 1
US-09-106-872A-22
; Sequence 22, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 22
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-22

Alignment Scores:
Pred. No.: 8.73e-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-09-106-872A-22 (1-157)


```
QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCTCTCTGCTGCCACGCATCTGCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGAGACAGAAGATGCCAGAGCCAGCTCAGAGGGCGGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAenLeuArg 40
QY 122 CCTCTGCGAGCAACATCTTCATGCAGAGATCCACGCTGACGAGGATTCATATGAACGGAC 181
Db 41 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATCATCGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGACCAAGAGAGGTTTGCATGATGAGCTGAAGCTTTCAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAenGluPheGluAenAenGlnArg 100
QY 302 TGCATGTGCGAGGCATTCACACAGATCATGGAGAACACAGAGCGATAGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAenGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAAGGCTTGCAGCTTGAGCTGCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157

RESULT 2
US-09-191-593-10
; Sequence 10, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M,
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
```

```
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; STRANDEDNESS: not applicable
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; DESCRIPTION: identified as Ara h II cDNA clone
; DESCRIPTION: derived amino acid sequence
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE: No. 6835824 applicable
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE: seed
; TISSUE TYPE: seed cDNA
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLS: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-10

Alignment Scores:
Pred. No.: 8,73e-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.56% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-09-191-593-10 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTTCTCTCTGCTGCCACGCATCTGCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGAGACAGAAGATGCCAGAGCCAGCTCAGAGGGCGGAACCTGAGG 121
```


MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS 6.2
SOFTWARE: Wordperfect 6.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/191,593
FILING DATE: 13 NOVEMBER 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/717,933
FILING DATE: 23 SEPTEMBER 1996
APPLICATION NUMBER: US 07/998,377
FILING DATE: 30 DECEMBER 1992
APPLICATION NUMBER: US 08/158,704
FILING DATE: 29 NOVEMBER 1993
APPLICATION NUMBER: US 60/009,455
FILING DATE: 29 DECEMBER 1995
APPLICATION NUMBER: US 08/610,424
FILING DATE: 04 MARCH 1996
ATTORNEY/AGENT INFORMATION:
NAME: ALEXANDER, DANIEL R
REGISTRATION NUMBER: 32,604
REFERENCE/DOCKET NUMBER: ARK00895601B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (501) 582-9111
TELEFAX: (501) 521-4931
TELEX: No. 6835824 applicable
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
STANDEDNESS: No. 6835824 applicable
TOPOLOGY: unknown
MOLECULE TYPE: glycoprotein
DESCRIPTION: identified as Ara h 2 p38 deduced
from nucleotide sequence reading frame.
HYPOTHETICAL: No
ANTI-SENSE: No. 6835824 applicable
FRAGMENT TYPE: No. 6835824 applicable
ORIGINAL SOURCE:
ORGANISM: Arachis hypogaea
STRAIN: Florunner
INDIVIDUAL ISOLATE: Ara h II
DEVELOPMENTAL STAGE:
HAPLOTYPE: No. 6835824 applicable
TISSUE TYPE:
CELL TYPE: No. 6835824 applicable
CELL LINE: No. 6835824 applicable
ORGANELLE: No. 6835824 applicable
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: No. 6835824 applicable
MAP POSITION: No. 6835824 applicable
UNITS: No. 6835824 applicable
FEATURE:
NAME/KEY: final check
LOCATION: 4..480
IDENTIFICATION METHOD: By agreement with
IDENTIFICATION METHOD: protein information and established
IDENTIFICATION METHOD: consensus sequence
OTHER INFORMATION: Seed storage protein and
allergen
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:

DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-21
Alignment Scores:
Pred. No.: 4,19e-84 Length: 157
Score: 827.00 Matches: 155
Percent Similarity: 98.73% Conservative: 0
Best Local Similarity: 98.73% Mismatches: 2
Query Match: 65.38% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-21 (1-157)
QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTTCCTCCCTCCACGCATCTCGAGGCGAG 61
DB 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGG 121
DB 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTCGAGCAACATCTCATGCAGAGATGCCAACGTGACGAGGATTTCATATGAACGGGAC 181
DB 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
DB 61 ProSerTyrProSerGlnAspProTyrSerProTyrSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAACAACAAAGG 301
DB 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGCGATTGCAACAGATCATGAGAGAACCGAGAGCGATAGTTGCAGGGGAGG 361
DB 101 CysMetCysGluAlaLeuGlnGlnMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTTCCTCAACAGTGGCGCTTAGG 421
DB 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAACAGCTTGGCACTTGGACGTGCGAAAGTGGCGGCGAGAGACAGATAC 472
DB 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157
RESULT 5
US-08-618-911-4
Sequence 4, Application US/08618911
Patent No. 5850016
GENERAL INFORMATION:
APPLICANT: Jung, Rudolf
APPLICANT: Haetings, Craig
APPLICANT: Coughlan, Sean
APPLICANT: Hu, David
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
SEEDS
TITLE OF INVENTION: SEEDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: USA
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

/ REGISTRATION NUMBER: 32,604
/ REFERENCE/DOCKET NUMBER: ARK00895601B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (501) 582-9111
/ TELEFAX: (501) 521-4931
/ TELEX: NO. 6835824 applicable
/ INFORMATION FOR SEQ ID NO: 53:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 28 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: NO. 6835824 applicable
/ TOPOLOGY: unknown
/ MOLECULE TYPE: glycoprotein
/ DESCRIPTION: identified as derived N-terminal
/ HYPOTHETICAL: No. 6835824 applicable
/ ANTI-SENSE: NO. 6835824 applicable
/ FRAGMENT TYPE: N-terminal fragment
/ ORIGINAL SOURCE:
/ ORGANISM: Atachis hypogaea
/ STRAIN: Florunner
/ INDIVIDUAL ISOLATE: P38
/ DEVELOPMENTAL STAGE:
/ HAPLOTYPE: No. 6835824 applicable
/ TISSUE TYPE:
/ CELL TYPE: No. 6835824 applicable
/ CELL LINE: No. 6835824 applicable
/ ORGANELLE: No. 6835824 applicable
/ IMMEDIATE SOURCE:
/ LIBRARY:
/ CLONE: P38
/ POSITION IN GENOME:
/ CHROMOSOME/SEGMENT: No. 6835824 applicable
/ MAP POSITION: No. 6835824 applicable
/ UNITS: No. 6835824 applicable
/ FEATURE:
/ NAME/KEY:
/ LOCATION:
/ IDENTIFICATION METHOD: By agreement with
/ IDENTIFICATION METHOD: protein information and established
/ IDENTIFICATION METHOD: consensus sequence
/ OTHER INFORMATION: Seed storage protein and
/ OTHER INFORMATION: allergen
/ PUBLICATION INFORMATION:
/ AUTHORS:
/ TITLE:
/ JOURNAL:
/ VOLUME:
/ ISSUE:
/ PAGES:
/ DATE:
/ DOCUMENT NUMBER:
/ FILING DATE:
/ PUBLICATION DATE:
/ RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-53

Alignment Scores:
Pred. No.: 3.78e-09 Length: 28
Score: 156.00 Matches: 28
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.33% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-09-191-593-53 (1-28)

QY 59 CAGCAGTGGGAATCTCAAGAGAGACAGAAGATGCCAGAGCGAGCTCGAGAGGGCGAACCCTG 118
Db 1 GlnGlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeu 20

QY 119 AGGCCCTGCGAGCAACATCTCATG 142
|||||

Db 21 ArgProCysGluGlnHisLeuMet 28

RESULT 11
US-08-670-186-4
/ Sequence 4, Application US/08670186
/ Patent No. 5859343
/ GENERAL INFORMATION:
/ APPLICANT: SUN, SAMUEL S.M.
/ APPLICANT: XIONG, LIWEN
/ APPLICANT: HU, ZHONG
/ APPLICANT: CHEN, HANG
/ TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
/ NUMBER OF SEQUENCES: 10
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: MORRISON & FOERSTER
/ STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
/ CITY: WASHINGTON
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20006-1888
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/670,186
/ FILING DATE: 21-JUN-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MURASHIGE, KATE H.
/ REGISTRATION NUMBER: 29,959
/ REFERENCE/DOCKET NUMBER: 23461-20007.00
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 887-1500
/ TELEFAX: (202) 822-0168
/ TELEX: 90-4030 MRSNFOERSWSH
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 158 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-670-186-4

Alignment Scores:
Pred. No.: 7.17e-09 Length: 158
Score: 156.00 Matches: 46
Percent Similarity: 47.56% Conservative: 32
Best Local Similarity: 28.05% Mismatches: 52
Query Match: 12.33% Indels: 34
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x US-08-670-186-4 (1-158)

QY 2 CTCACCATAGTAGAGCCCTCGCCCTTTTCCTCCGTCGCCAGCATCTGCG---AGG 58
Db 5 IleLeuLeuLeuThrThrLeuAlaLeuPheValLeuLeuAlaAsnAlaSerIleTyArg 24
QY 59 CAGCAGTGGGAATCTC-----CAAGAGACAGAAGATGCCAGAGC 97
Db 25 ThrThrValGluLeuAspGluAspAsnAspGluAsnGlnProLeuCysArgArg 44
QY 98 CAGCTCGAGAGG---CGGAACCTGAGGCCCTCGAGCAACATCTCATGCAGAAAGATCCAA 154
Db 45 GlnPheGlnGlnHisGlnHisLeuArgAlaCysGlnArgTyIleArgArgAlaGln 64
QY 155 CQTGAC-----GAGGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTAC 208
Db 65 ArgGlyGlyLeuValAspGluLeuGluAspValGluGluAsnGluAsp----- 82
QY 209 AGCCCTAGTCCATATGATCGGAGAGCGCTGATCTCTCAGCACCAAGAGAGG----- 262
|||||


```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-670-186-2

Alignment Scores:
Pred. No.: 5,73e-07 Length: 155
Score: 139.00 Matches: 45
Percent Similarity: 45.51% Conservative: 26
Best Local Similarity: 28.85% Mismatches: 47
Query Match: 10.99% Indels: 38
DB: 2 Gaps: 9

US-10-728-323-2 (1-717) x US-08-670-186-2 (1-155)
QY 8 ATACTAGTACCTCGCCCTTTCTCTCTCGCCACGCACTCTCGAGG----- 58
Db 7 LeuPheAlaThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleGlnThrVal 26
QY 59 -----CAGCAGTGGGAACCTCCAAGGAGACAGAAGATGC 91
Db 27 ValGluValAspGluGluAspAsnGlnLeuTip-----ArgCys 40
QY 92 CAGAGCCAG---CTCGAGAGGGCGAACCTGAGGCCCTCGAGCAACATCTCATGCGAAG 148
Db 41 GlnArgGlnPheLeuGlnHisGlnArgLeuArgAlaCysGlnArgPheIleHisArgArg 60
QY 149 ATCCAA-----CGTGACGAGGATTCATATGAACGGACCCGTACAGCCCTAGTCAGGAT 202
Db 61 AlaGlnPheGlyGlnProAspGluLeuGluAspGluValGluAspAsnAspAsp 80
QY 203 CGGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGG 262
Db 81 GluAsnGlnPro-----ArgArgProAla-----LeuArgGln 91
QY 263 TGTTCGATGAGTGAACGAGTTTGAGAACCAACCAAGGTGCGATGTC----- 310
Db 92 CysCysAsnGlnLeuArgGlnValAsp-----ArgProCysValCysProValLeuArg 109
QY 311 GAGCGATTGCAACAGATCATGTGAGAACCCAGGCGATAGTTGTCAGGGGAGGCAACAGGAG 370
Db 110 GluAlaGlnGlnValLeuGlnArgGlnIle-----IleGlnGlyProGlnGlnLeu 127
QY 371 CAACAGTTCAAGAGGAGCTCAGAACTTCCTCAACAGTGGCGCTT 418
Db 128 ArgArgLeuPheAspAlaAlaArgAsnLeuProAsnIleCysAsnIle 143

RESULT 14
US-08-453-924-3
; Sequence 3, Application US/08453924
; Patent No. 5608152
; GENERAL INFORMATION:
; APPLICANT: Kridl, Jean C.
; TITLE OF INVENTION: Seed-Specific Transcriptional Regulation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; STREET: 2882 Sand Hill Road, Ste. 280
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,924
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/742,834
```

RESULT 15
US-09-191-593-58
; Sequence 58, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M.
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; APPLICANT: BANNON, Gary A
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6835824 applicable
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; DESCRIPTION: identified as 17.5 kD N-terminal
; DESCRIPTION: sequence of Ara h 2 p38 (gene sequence 20..47).
; HYPOTHETICAL: No
; ANTI-SENSE: No. 6835824 applicable
; FRAGMENT TYPE: amino terminus
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE: P38
; DEVELOPMENTAL STAGE:
; HAPLOTYPE: No. 6835824 applicable
; TISSUE TYPE:
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLE: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY: P38
; CLONE: P38

; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-58
Alignment Scores:
Pred. No.: 3.95e-07 Length: 28
Score: 138.00 Matches: 26
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 2
Query Match: 10.91% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-09-191-593-58 (1-28)
QY 59 CAGCAGTGGGAAGTCCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
Db 1 GlnGlnTrpGluLeuGlnGlyAspArgArgGlnSerGlnLeuGluArgAlaAsnLeu 20
QY 119 AGGCCCTCGAGCAACATCTCATG 142
Db 21 ArgProCysGluGlnHisLysMet 28
Search completed: August 24, 2005, 10:09:12
Job time : 15.417 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 24, 2005, 10:05:17 ; Search time 46.2283 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1265
Sequence: 1 gctcaccatactagtagcc.....taaaagatcatgttttgtt 717

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DRV=xlh
-O=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124357_29323/app.query.fasta_1.4757
-DB=Published Applications AA -QFMT=fastan -SURF=n2p.rapp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10728323 @CIGN 1 1 221 @runat_23082005_124357_29323
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	842	66.6	157	14	US-10-228-806-4	Sequence 4, Appl
2	842	66.6	157	15	US-10-100-303A-63	Sequence 63, Appl
3	838.5	66.3	207	17	US-10-958-324-2	Sequence 2, Appl
4	835	66.0	160	17	US-10-899-551-4	Sequence 4, Appl
5	819	64.7	156	15	US-10-243-871-53	Sequence 53, Appl
6	819	64.7	156	15	US-10-253-286-53	Sequence 53, Appl
7	771	60.9	166	15	US-10-100-303A-82	Sequence 82, Appl
8	760	60.1	167	17	US-10-899-551-56	Sequence 56, Appl
9	699	55.3	166	9	US-09-731-221-77	Sequence 77, Appl
10	463	36.6	83	15	US-10-100-303A-81	Sequence 81, Appl
11	243	19.2	158	15	US-10-302-633-2	Sequence 2, Appl
12	243	19.2	158	15	US-10-424-599-169225	Sequence 169225,
13	241.5	19.1	155	15	US-10-424-599-264372	Sequence 264372,
14	177.5	14.0	168	14	US-10-165-289A-2	Sequence 2, Appl
15	149.5	11.8	167	16	US-10-382-086-29	Sequence 29, Appl
16	144	11.4	169	14	US-10-165-289A-4	Sequence 4, Appl
17	140.5	11.1	165	16	US-10-425-115-190971	Sequence 190971,
18	139	11.0	184	11	US-09-782-130-3	Sequence 3, Appl
19	139	11.0	184	14	US-10-100-121-48	Sequence 48, Appl
20	129.5	10.2	150	15	US-10-442-174A-1	Sequence 1, Appl
21	127.5	10.1	151	15	US-10-424-599-167336	Sequence 167336,
22	127	10.0	122	14	US-10-228-806-67	Sequence 67, Appl
23	126	10.0	102	15	US-10-424-599-219338	Sequence 219338,
24	124	9.8	148	16	US-10-425-115-219350	Sequence 219350,
25	119	9.4	20	9	US-09-731-221-66	Sequence 66, Appl
26	118	9.2	261	15	US-10-291-265-715	Sequence 715, App
27	117	9.2	20	9	US-09-731-221-63	Sequence 63, Appl
28	116	9.2	20	9	US-09-731-221-60	Sequence 60, Appl
29	115	9.1	20	9	US-09-731-221-61	Sequence 61, Appl
30	113	8.9	20	9	US-09-731-221-65	Sequence 65, Appl
31	112.5	8.9	101	15	US-10-424-599-275916	Sequence 275916,
32	112	8.9	20	9	US-09-731-221-54	Sequence 54, Appl
33	112	8.9	20	9	US-09-731-221-59	Sequence 59, Appl
34	112	8.9	20	9	US-09-731-221-62	Sequence 62, Appl
35	112	8.9	20	9	US-09-731-221-64	Sequence 64, Appl
36	111	8.8	20	9	US-09-731-221-67	Sequence 67, Appl
37	111	8.8	20	9	US-09-731-221-74	Sequence 74, Appl
38	111	8.8	20	9	US-09-731-221-75	Sequence 75, Appl
39	111	8.8	20	9	US-09-731-221-76	Sequence 76, Appl
40	110	8.7	25	15	US-10-245-871-60	Sequence 60, Appl
41	110	8.7	25	15	US-10-253-286-60	Sequence 60, Appl
42	109	8.6	20	15	US-10-245-871-55	Sequence 55, Appl
43	109	8.6	20	15	US-10-253-286-55	Sequence 55, Appl
44	109	8.6	149	18	US-10-481-032A-244	Sequence 244, App
45	108	8.5	20	9	US-09-731-221-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-10-228-806-4
; Sequence 4, Application US/10228806
; Publication NO. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-4

Alignment Scores:

Pred. No.:	1.15e-79	Length:	157
Score:	842.00	Matches:	157
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	66.56%	Indels:	0
DB:	14	Gaps:	0
US-10-728-323-2 (1-717) x US-10-228-806-4 (1-157)			
QY	2	CTCACCACTACTAGTACGCTCGCCCTTTCTCCTCGCTGCCACGCATCTCGAGGCGAG	61
Db	1	LeuthrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	20
QY	62	CAGTGGGAACCTCAAGGAGACAGAGAATGCCAGAGCGAGCTCGAGGGCGCACTCTGAGG	121
Db	21	GlnTrpLeuGlnGlnArgArgCysGlnSerGlnLeuGluArgAlaSerAlaArg	40
QY	122	CCTCGGAGCAACATCTCATGCAAGATGCCAGAGCTCAACCGTGCAGAGATTTCATATGAACGGAC	181
Db	41	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	60
QY	182	CGGTACAGCCCTAGTACGAGTCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA	241
Db	61	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	80
QY	242	TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGCAACAGTTCGAGAACCAACAAAGG	301
Db	81	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	100
QY	302	TGCATGTGCGAGGCATTCGAACAGATCATGGAAGACCGAGAGCGATAGTTGCAGGGGAGG	361
Db	41	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	60
QY	182	CGGTACAGCCCTAGTACGAGTCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA	241
Db	61	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	80
QY	242	TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGCAACAGTTCGAGAACCAACAAAGG	301
Db	81	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	100
QY	302	TGCATGTGCGAGGCATTCGAACAGATCATGGAAGACCGAGAGCGATAGTTGCAGGGGAGG	361
Db	101	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	120
QY	362	CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGCAGGCTTAGG	421
Db	121	GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	140
QY	422	GCACCACAGCGTTGCGACTTGGAACGTGCGAAAGTGGCGGCAGAGACAGATAC	472
Db	141	AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr	157
RESULT 2			
US-10-100-303A-63			
; Sequence 63, Application US/10100303A			
; Publication No. US20030202980A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, et al.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction			
; FILE REFERENCE: 2002834-0166			
; CURRENT APPLICATION NUMBER: US/10/100,303A			
; PRIOR FILING DATE: 2002-03-18			
; NUMBER OF SEQ ID NOS: 138			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 63			
; LENGTH: 157			
; TYPE: PRT			
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2			
US-10-100-303A-63			
Alignment Scores:			
Pred. No.:	1.15e-79	Length:	157
Score:	842.00	Matches:	157
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	66.56%	Indels:	0
DB:	15	Gaps:	0
US-10-728-323-2 (1-717) x US-10-100-303A-63 (1-157)			
QY	2	CTCACCACTACTAGTACGCTCGCCCTTTCTCCTCGCTGCCACGCATCTCGAGGCGAG	61
Db	1	LeuthrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	20
US-10-958-324-2			
; Sequence 2, Application US/10958324			
; Publication No. US20050114924A1			
; GENERAL INFORMATION:			
; APPLICANT: DODO, HORTENSE W.			
; APPLICANT: ARNTZEN, CHARLES J.			
; APPLICANT: KONAN, KOFFI N'DA			
; APPLICANT: VIOQUEZ, OLGA			
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN			
; FILE REFERENCE: 072121/0104			
; CURRENT APPLICATION NUMBER: US/10/958,324			
; CURRENT FILING DATE: 2004-10-06			
; PRIOR APPLICATION NUMBER: US/09/715,036			
; PRIOR FILING DATE: 2000-11-20			
; PRIOR APPLICATION NUMBER: 60/167,255			
; PRIOR FILING DATE: 1999-11-19			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 2			
; LENGTH: 207			
; TYPE: PRT			
; ORGANISM: Arachis hypogaea			
US-10-958-324-2			
Alignment Scores:			
Pred. No.:	2.89e-79	Length:	207
Score:	838.50	Matches:	162
Percent Similarity:	90.27%	Conservative:	5
Best Local Similarity:	87.57%	Mismatches:	17
Query Match:	66.28%	Indels:	1
DB:	17	Gaps:	1
US-10-728-323-2 (1-717) x US-10-958-324-2 (1-207)			
QY	2	CTCACCACTACTAGTACGCTCGCCCTTTCTCCTCGCTGCCACGCATCTCGAGGCGAG	61
Db	4	LeuthrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	23
QY	62	CAGTGGGAACCTCAAGGAGACAGAGAATGCCAGAGCGAGCTCGAGGGCGCAACCTGAGG	121
Db	24	GlnTrpLeuGlnGlnArgArgCysGlnSerGlnLeuGluArgAlaSerAlaArg	43

```
QY 122 CCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTGAGACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCCAGGACATTGCAACACATCATGTGAGAACACAGAGCGATAGTGTGCGAGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAGAGGTTCCGACTTGGAGCTGAAAGTGGCGGCGAGACAGATCTAAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgArgProArgIleProPro 163
QY 482 ATCTCAAAAAAGAAAGAAAGAAAGAAATAGCTTATATATAAGC---TATTATCTA 538
Db 164 IleuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183
QY 539 TGGTTATGTTAGTT 553
Db 184 AlaMetCysLeuLeu 188

RESULT 4
US-10-899-551-4
; Sequence 4, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 160
; TYPE: PRT
; ORGANISM: species Arachis hypogaea
US-10-899-551-4

Alignment Scores:
Pred. No.: 6,31e-79 Length: 160
Score: 835.00 Matches: 156
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.01% Indels: 0
DB: 17 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-4 (1-160)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCGATCTCGAGGACG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCCAAGGAGACAGAAAGATGCCAGAGCGAGCTCGAGAGGCGCAACTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43

US-10-728-323-2 (1-717) x US-10-245-871-53 (1-156)

QY 2 CTCACCATCTAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCGATCTCGAGGACG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCCAAGGAGACAGAAAGATGCCAGAGCGAGCTCGAGAGGCGCAACTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43

QY 122 CCCTGCGAGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGATTGAGACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCCAGGACATTGCAACACATCATGTGAGAACACAGAGCGATAGTGTGCGAGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTGGCGGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAGAGGTTCCGACTTGGAGCTGAAAGTGGCGGCGAGACAGATCTAAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgArgProArgIleProPro 159

RESULT 5
US-10-245-871-53
; Sequence 53, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-53

Alignment Scores:
Pred. No.: 3,06e-77 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
DB: 15 Gaps: 0
```


Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTCCAGGCGATTCGACAGATCATCGGAGAACAGAGCGATAGGTTCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGCAGCCTTAGG 421
Db 124 GlnGlnGlnGlnGlnPheGlyAsnGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 6
US-10-253-286-53
; Sequence 53, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PROT
; ORGANISM: Arachis hypogaea
US-10-253-286-53
Alignment Scores:
Pred. No.: 3,06e-77 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.74% Indels: 0
Gaps: 0
US-10-728-323-2 (1-717) x US-10-253-286-53 (1-156)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGCTGCCACGCGCATCTGCGAGGCAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCAAGAGACAGAAATGCGAGCGAGCTCGAGAGGGCGGACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGGAGCAACATCTTCACAGAGATCCACGTCGACGAGATTATATGAACGGGAC 181
Db 44 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGlnAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGGCTTGCATGACCTGAACAGATTGAGAAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTCCAGGCGATTCGACAGATCATCGGAGAACAGAGCGATAGGTTCAGGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGCAGCCTTAGG 421

Db 124 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACACACAGCGTTGCGACTTGGACGTGCGAAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 7
US-10-100-303A-82
; Sequence 82, Application US/10100303A
; Publication No. US20030202960A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 166
; TYPE: PROT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-82
Alignment Scores:
Pred. No.: 3,61e-72 Length: 166
Score: 771.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.95% Indels: 0
Gaps: 0
US-10-728-323-2 (1-717) x US-10-100-303A-82 (1-166)
QY 50 TCTGCGAGCGACGAGTGGGAATCCAAAGGAGACAGAAGATGCCAGAGCGAGCTCGAGAGG 109
Db 15 SerAlaArgGlnGlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArg 34
QY 110 GCGAACCTGAGCCCTCGGAGCAACATCTCATGCGAGAGATCCAAAGTGCAGGATTTCA 169
Db 35 AlaAsnLeuArgProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGluAspSer 54
QY 170 TATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGG 229
Db 55 TyrGluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 74
QY 230 AGAGCGCTGGATTCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGlu 94
QY 290 ACAACCAAGGTGCATGTGCGAGGCATTCGACAGATTCGACAGATTCGACAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
QY 350 TTGCGAGGGAGGCAACAGGAGCAACAGGAGCAACAGTTCGAAGAGGAGCTCAGGAACCTTGCTCAACAG 409
Db 115 LeuGlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGln 134
QY 410 TCGCGCCTTAGGGCCACCACAGCGTTGCGACTTGGACGTGGAAGTGGCGGCGAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 154
QY 470 TAC 472
Db 155 Tyr 155
RESULT 8
US-10-899-551-56
; Sequence 56, Application US/10899551
; Publication No. US2005006394A1
; GENERAL INFORMATION:

APPLICANT: Caplan, Michael J.
APPLICANT: Burks, A. Wesley
APPLICANT: Sampson, Hugh A.
APPLICANT: Howard, Sasin B.
APPLICANT: Bottomly, Kim H.
TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
FILE REFERENCE: 2002834-0233
CURRENT APPLICATION NUMBER: US/10/899,551
CURRENT FILING DATE: 2004-07-26
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.2
SEQ ID NO 56
LENGTH: 167
TYPE: PRT
ORGANISM: species Arachis hypogaea
US-10-899-551-56

Alignment Scores:
Pred. No.: 5,23e-71 Length: 167
Score: 760.00 Matches: 139
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.08% Indels: 0
DB: 17 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-56 (1-167)

Qy 53 GCGAGGAGCAGTGGGAATCTCCAGAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGCG 112
Db 17 AlaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAla 36
Qy 113 AACCTGAGGCCCTCGAGCAACATCTCATGACAGAAATCCACCTGACGAGATTCATAT 172
Db 37 AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyr 56
Qy 173 GAACGGACCCGTACAGCCCTAGTCAGATCCGACAGCCCTAGTCCATATGATCGAGA 232
Db 57 GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArg 76
Qy 233 GGCCTGGATCCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGGAACGAGTTTGAAAC 292
Db 77 GlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsn 96
Qy 293 AACCAAGGTGCATGTGCGAGGCATTCACAGATCATGAGAACCCAGAGCGATAGTTG 352
Db 97 AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeu 116
Qy 353 CAGGGAGGCAACAGGAGCAACAGTTCAGAGGAGGCTCAGGAAGTGGCGGCAGAGACAGA 412
Db 117 GlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCys 136
Qy 413 GGCCTTAGGGCACCACAGCGTTCGACCTTGACCTCGAAAGTGGCGGCAGAGACAGA 469
Db 137 GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 155

RESULT 9

US-09-731-221-77
Sequence 77, Application US/09731221
Patent No. US2002018778A1
GENERAL INFORMATION:
APPLICANT: Caplan, Michael
TITLE OF INVENTION: Passive Desensitization
FILE REFERENCE: 2002834-0103
CURRENT APPLICATION NUMBER: US/09/731,221
CURRENT FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 79
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 77
LENGTH: 166
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Arachis

OTHER INFORMATION: Hypogaea
US-09-731-221-77

Alignment Scores:
Pred. No.: 1,42e-64 Length: 166
Score: 699.00 Matches: 131
Percent Similarity: 92.91% Conservatives: 0
Best Local Similarity: 92.91% Mismatches: 10
Query Match: 55.26% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x US-09-731-221-77 (1-166)

Qy 50 TCTGCGAGCAGCAGTGGGAATCTCCAGAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAG 109
Db 15 SerAlaArgGlnGlnAlaGluLeuGlnGlyAspArgCysGlnSerGlnLeuAlaArg 34
Qy 110 GCGAACCTGAGGCCCTCGAGCAACATCTCATGACAGAAATCCAACTGACGAGGATTCA 169
Db 35 AlaAsnLeuArgAlaCysGluAlaHisLeuMetGlnLysIleGlnAlaAspGluAspSer 54
Qy 170 TATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGG 229
Db 55 TyrGluArgAlaProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArg 74
Qy 230 AGAGGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTGAG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGlu 94
Qy 290 AACCAACCAAGGTGCATGTGCGAGGCATTCGCAACAGATCATGAGAACCCAGAGCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
Qy 350 TTGACAGGAGGAGCAACAGGAGCAACAGTTCAGAGGAGGCTCAGGAAGTTCCTCAACAG 409
Db 115 LeuGlnGlyAlaGlnGlnGlnGlnPheLysArgGluAlaArgAsnLeuProGlnGln 134
Qy 410 TGGCGCTTAGGGCACCACAGCCTTGGACCTTGGACCTCGAAAGTGGCGGCAGAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArg 154
Qy 470 TAC 472
Db 155 Tyr 155

RESULT 10

US-10-100-303A-81
Sequence 81, Application US/10100303A
Publication No. US20030202980A1
GENERAL INFORMATION:
APPLICANT: Caplan, et al.
TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
FILE REFERENCE: 2002834-0166
CURRENT APPLICATION NUMBER: US/10/100,303A
CURRENT FILING DATE: 2002-03-18
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 81
LENGTH: 83
TYPE: PRT
ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-81

Alignment Scores:
Pred. No.: 9,37e-40 Length: 83
Score: 463.00 Matches: 83
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 36.60% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-2 (1-717) x US-10-100-303A-81 (1-83)

Qy	68	GAACTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTTGAGGCCCTGC	127
Db	1	GLuLeuGLnGLyAspArgargCysGlnSerGlnLeuGluAraGlaAenLeuAraProCys	20
Qy	128	GAGCAACATCTCATGCAGAGATCCAACGTGACGAGGATTTCATATGAACGGACCCGTAC	187
Db	21	GLuGLNHisLeuMetGlnLysIleGlnArgaspGluaspSerTyrGluAra-gaspProTyr	40
Qy	188	AGCCCTAGTCAGGATCCGTACACGCCCTAGTTCATATGATCGAGAGCGCTGGATCCTCT	247
Db	41	SerProSerGlnAspProTyr-SerProSerProTyrAspArgagrgGlyAlaGlySerSer	60
Qy	248	CAGCACCAAGAGAGGTGTTCGAATCAGCTGACGAGTTTGAGAACCAACAAGGTGCATG	307
Db	61	GlnHisGlnGluArgCysCysaaengluLeuaengluPheGluAenAenGlnAra-gcysMet	80
Qy	308	TGCAGGCCA 316	
Db	81	Cysgluala 83	
RESULT 11			
US-10-302-633-2			
; Sequence 2, Application US/10302633			
; Publication No. US20030229038A1			
; GENERAL INFORMATION:			
; APPLICANT: de Lumen, Benito O.			
; Galvez, Alfredo F.			
; TITLE OF INVENTION: Lunasin Peptides			
; NUMBER OF SEQUENCES: 3			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP			
; STREET: 75 DENISE DRIVE			
; CITY: HILLSBOROUGH			
; STATE: CALIFORNIA			
; COUNTRY: USA			
; ZIP: 94010			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patentin Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/302,633			
; FILING DATE: 22-No. US20030229038A1-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/09/531,727			
; FILING DATE: 21-Mar-2000			
; APPLICATION NUMBER: 08/938,675			
; FILING DATE: <Unknown>			
; ATTORNEY/AGENT INFORMATION:			
; NAME: OSMAN, RICHARD A			
; REGISTRATION NUMBER: 36,627			
; REFERENCE/DOCKET NUMBER: B98-003			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (650) 343-4341			
; TELEFAX: (650) 343-4342			
; INFORMATION FOR SEQ ID NO: 2:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 158 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: peptide			
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:			
US-10-302-633-2			
Alignment Scores:			
Pred. No.:	1,828-16	Length:	158
Score:	243.00	Matches:	56
Percent Similarity:	56.4%	Conservative:	31
Best Local Similarity:	36.36%	Mismatches:	55

[illegible]

```

RESULT 13
US-10-424-599-264372
; Sequence 264372, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264372
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80749C.1.pep
US-10-424-599-264372

```

Alignment Scores:					
Pred. DO.:	2.6e-16	Length:	155		
Score:	241.50	Matches:	58		
Percent Similarity:	57.11%	Conservative:	30		
Best Local Similarity:	37.66%	Mismatches:	53		
Query Match:	19.09%	Indels:	13		
DB:	15	Gaps:	7		
US-10-728-323-2 (1-717)	x	US-10-424-599-264372 (1-155)			

QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTTCTCTCGTCCGCCACGCATCTCGCAGGCAG	61
Db	4	LeuThrIleLeuLeuIleAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer	21
QY	62	CAGTGGGAATCTCCAGGAGACAGAGATCCAGACCCAGCTCGAGCGCGCAACTCTAGG	12
Db	22	LysTrpGlnGlnHisGlnGlnGlnSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn	41
QY	122	CCCTGGCAGCAATCTCATGCAGAGATCCAA-----CGTGACGAGGATTCATAT	177
Db	42	ProCysGlu---HisIleMetGluLysIleGlnAlaGlyArgArgGlyGluAsnGlySer	60

Qy	173	GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA	232
Db	61	AspGluAspHisIleuLeuileArgThrMetPro---GlyArgIleAsnTrpIleArgLys	79
Qy	233	GGCGCTGGATCCTCTCAG-----CACCAAGAGAGAGTGTCGAATGAGCTGAAC	280
Db	80	LysGluGluLysGluGluGluGluGluGluGlyHisMetGlnLysCysSerGluMetSer	99
Qy	281	GAGTTTGAGAAACAACAAAGGTGCATGTGCGAGGCATTGCCAACATCATGGAGAACCCAG	340
Db	100	GluLeuLys---SerProIleCysGlnCysLysAlaLeuGlnLysIleMetAspAsnGln	118
Qy	341	AGCATAGCTGTTCAGGGGAGCAACAGCAGGACCAACAGTTTCAACAGGAGGAGCTCAGGAACCTTC	400
Db	119	SerGluGlnLeuGluGluLys---GluLysGlnMetGluArgGluLeuMetAsnLeu	137
Qy	401	CCTCAACAGTCGGCGCTTAGGGCCACCAAGCGTTGCGACTTG	442
Db	138	AlaIleArgCysArgLeuGluGlyProMetIleGlyCysAspLeu	151

```

RESULT 14
US-10-165-289A-2
; Sequence 2, Application US/10165289A
; Publication No. US20030159174A1
; GENERAL INFORMATION:
; APPLICANT: QIU, Xiao
; APPLICANT: TRUSKA, Martin
; APPLICANT: HU, Zhiyuan
; TITLE OF INVENTION: Flax (Linum usitatissimum L.)
; TITLE OF INVENTION: Seed-Specific Promoters
; FILE REFERENCE: BNZ-005
; CURRENT APPLICATION NUMBER: US/10/165,289A
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/295823
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Linum usitatissimum
US-10-165-289A-2

```

Alignment Scores:		
Pred. No.:	1.5e-09	Length:
Score:	177.50	Matches:
Percent Similarity:	45.71%	Conservative:
Best Local Similarity:	30.86%	Mismatches:
Query Match:	14.03%	Indels:
DB:	14	Gaps:
		9

US-10-728-323-2 (1-717) x US-10-165-289A-2 (1-168)

Qy	2	CTCAC	ATACTAGT	AGCCCTGCC	-----	CTTTT	CCTCTCGCTG	CCGCCACGCATCT	52
Db	4	LeuMetSer	LeuAlaValAla	ThrAlaPhe	LeuPheLeu	ValValAsp	AlaSer	23	
Qy	53	GCAGG	ACACAGTGGG	AACTCCA	AGGAGAC	-----	-----	82	
Db	24	ValArgThr	ThrVal	LeuLeuAsp	GluAspThr	ArgGlnGly	ArgGlyGlyGlnGlyGly	43	
Qy	83	-----	AGAAGATGC	CAGAGCCAGCTC	-----	GAGAGGCGCAAC	CTCGAGCCCTGC	12	
Db	44	GlnGlyGln	GlnGlnCys	gluLysGln	LeuGlnAsp	TyrLeuArgSerCys	63		
Qy	128	GAGCA	ACATCTCAT	CAGAGATCA	CAACGTGAC	GAGGATTCATATGA	ACGGGACCCGTAC	18	
Db	64	GlnPheLeu	TrpGluLys	ValGlnGly	ArgSerTyr	-----	-----	78	
Qy	188	AGCCCTAGT	CAGGATCC	GTACAGCCCTAGT	CCATATGATCGG	AGAGCGCGCTGGA	-----	241	
Db	79	-----	-----	-----	-----	-----	-----	87	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:17:40 ; Search time 4512.08 Seconds
(without alignments)
16366.227 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcgcagcaaccggaggagaa.....ctccgaggctgtgcttaa 1524

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sv.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1524	100.0	1524	6	BD107900	Methods a
2	1524	100.0	1524	8	AF093541	Arachis h
3	1509.6	99.1	1524	6	AX155334	Sequence
4	1391.2	91.3	1853	6	AX148741	Sequence
5	1391.2	91.3	1853	8	AF086821	Arachis h
6	1373.2	90.1	1590	8	AF125192	Arachis h
7	1329.4	87.2	1901	8	AY722686	Arachis h
8	1328	87.1	1934	8	AY722685	Arachis h
9	1321.6	86.7	1886	8	AY439332	Arachis h
10	854.2	56.0	1533	8	AF510854	Arachis h
11	834.6	54.8	3825	8	AF510854	Arachis h
12	694.8	45.6	1786	8	AY722687	Arachis h
13	519.2	34.1	657	8	AF487543	Arachis h
14	517.4	34.0	1446	6	BD175913	Regulatio
15	517.4	34.0	1446	6	AR202578	Sequence
16	517.4	34.0	1446	6	AR486465	Sequence
17	517.4	34.0	1446	6	AX839958	Sequence
18	517.4	34.0	1446	6	AX840483	Sequence
19	515.8	33.8	1696	8	AB030494	Glycine m

20	513	33.7	1488	6	BD175911	Regulatio
21	513	33.7	1488	6	AR202576	Sequence
22	513	33.7	1488	6	AR486463	Sequence
23	513	33.7	1488	6	AX839950	Sequence
24	513	33.7	1488	6	AX840475	Sequence
25	513	33.7	1550	8	AB113349	Glycine m
26	513	33.7	1743	6	E32414	Transgenic
27	513	33.7	1743	6	AR343131	Sequence
28	513	33.7	1743	8	SOXGLYBSU	Soybean gly
29	505.6	33.2	1708	8	SOYGA2B1A	Glycine max
30	505	33.1	1746	6	E02463	cdNA encodi
31	505	33.1	1746	6	AX252313	Sequence
32	505	33.1	1746	8	GMGLY1A	Soybean mRN
33	504	33.1	1458	6	BD175912	Regulatio
34	504	33.1	1458	6	AR202577	Sequence
35	504	33.1	1458	6	AR486464	Sequence
36	504	33.1	1458	6	AX839952	Sequence
37	504	33.1	1458	6	AX840477	Sequence
38	504	33.1	1639	8	AB113350	Glycine m
39	502.4	33.0	1712	6	E02462	cdNA encodi
40	502.4	33.0	1712	8	GMGLY1	Glycine max
41	502.2	33.0	1710	8	AB030495	Glycine m
42	477.4	31.3	1729	8	PSA132614	Pisum sat
43	431.8	28.3	1664	8	VSLEGUMA	Vicia mR
44	426.2	28.0	1666	8	VFLEA2	Vicia faba
45	421.8	27.7	1850	8	VNPROLEGA	V.narbonens

ALIGNMENTS

RESULT 1
BD107900
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD107900
Methods and reagents for decreasing allergic reactions.
BD107900
BD107900.1 GI:23202718
JP 2002501748-A/3.
unidentified
unclassified.
1 (bases 1 to 1524)
Sosin, H., Banon, G.A., Jr, W.A.B. and Samphthon, H.A.
Methods and reagents for decreasing allergic reactions
Patent: JP 2002501748-A 3 22-JAN-2002;
UNIVERSITY OF ARKANSAS, MOUNT SINAI SCHOOL OF MEDICINE OF THE CITY
UNIVERSITY OF NEW YORK, HOWARD SOSIN
OS Arachis L. (Peanut)
PN JP 2002501748-A/3
PD 22-JAN-2002
PF 29-JAN-1999 JP 2000529437
PR 31-JAN-1998 US 60/073283,13-FEB-1998 US 60/074590 PR
13-FEB-1998 US 60/074624,13-FEB-1998 US 60/074633 PR
27-AUG-1998 US 09/141220
PI HAWADO SOSIN, GARY A BANON, WESLEY A BIRX JR, HYU A SAMPTHON PC
C12N15/09,A01H5/00,A01K67/027,A61K39/35,C07K14/37,C07K14/415, PC
C07K14/435,
PC C07K14/47,C12N1/15,C12N1/19,C12N1/21,C12N5/00,C12N15/00,C12N5/00
PC 00
CC Methods and reagents for decreasing allergic reactions PH
Key source Location/Qualifiers
FT
FT
FT

FEATURES
source

ORIGIN

Query Match 100.0%; Score 1524; DB 6; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGCAGCAACCGGAGGAGAACGGGTGCCAGTTCACAGCGCTCAATCGCAGAGACCTGAC 60
Db |||
1 CGCAGCAACCGGAGGAGAACGGGTGCCAGTTCACAGCGCTCAATCGCAGAGACCTGAC 60
QY 61 AATCGCATTTGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACGAGAGTTC 120
Db |||
61 AATCGCATTTGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACACGAGAGTTC 120
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGTTAGTTCCTCGCGCAACGCCCTTCGTAGGCT 180
Db |||
121 GAATGCGCGCGCTCGCCCTCTCTCGTTAGTTCCTCGCGCAACGCCCTTCGTAGGCT 180
QY 181 TTCTACTTCCAATGCTCCCAAGGAGATCTTCATCAGCAAGGAAGGGGATACATTTGGGTTG 240
Db |||
181 TTCTACTTCCAATGCTCCCAAGGAGATCTTCATCAGCAAGGAAGGGGATACATTTGGGTTG 240
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAAAGCCCTCACACAAGGTTCGTGATCTCAG 300
Db |||
241 ATATTCCCTGGTGTCTTAGACACTATGAAAGCCCTCACACAAGGTTCGTGATCTCAG 300
QY 301 TCCCAAGACCCCAAGAGTCTCCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db |||
301 TCCCAAGACCCCAAGAGTCTCCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
QY 361 CACCAGAAGGTGCACCGTTTTCGATGAGGAGTCTCAITTCAGTTCCTCCACCGGTGTGCT 420
Db |||
361 CACCAGAAGGTGCACCGTTTTCGATGAGGAGTCTCAITTCAGTTCCTCCACCGGTGTGCT 420
QY 421 TTCTGGCTCTAACAGCACGACACTGATGTTGTTGCTGTTTCTTACTGACACCAAC 480
Db |||
421 TTCTGGCTCTAACAGCACGACACTGATGTTGTTGCTGTTTCTTACTGACACCAAC 480
QY 481 AACAAACGACACCGCTTGATCAGTTCCCGAGGAGATTCAATTTGCTGGCAACGAGAG 540
Db |||
481 AACAAACGACACCGCTTGATCAGTTCCCGAGGAGATTCAATTTGCTGGCAACGAGAG 540
QY 541 CAAGAGTTCCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
Db |||
541 CAAGAGTTCCTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
QY 601 AGCCCATACAGCCCGCAAGTACGCTAGACAAAGAGAGCGTGAAATTTAGCCCTCAGGA 660
Db |||
601 AGCCCATACAGCCCGCAAGTACGCTAGACAAAGAGAGCGTGAAATTTAGCCCTCAGGA 660
QY 661 CAGCAGCGCGCAGAGAAACGAGCAGACGAAGAAAGAAACGAAGTGGAAACATCTTC 720
Db |||
661 CAGCAGCGCGCAGAGAAACGAGCAGACGAAGAAAGAAACGAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTCAACGCGGAGTTCTTGGAAACGAGCCTTCCAGGTTGACGACAGACAGATAGTG 780
Db |||
721 AGCGGCTTCAACGCGGAGTTCTTGGAAACGAGCCTTCCAGGTTGACGACAGACAGATAGTG 780
QY 781 CAAAACTTAAGCGCAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACAGTGGGGA 840
Db |||
781 CAAAACTTAAGCGCAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACAGTGGGGA 840
QY 841 GGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGAATACGAT 900
Db |||
841 GGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
Db |||
901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
QY 961 GGGAAATGGTATGAAGAGAGATCTGCAACGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
Db |||
961 GGGAAATGGTATGAAGAGAGATCTGCAACGCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC 1080
Db |||
1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAAC 1080
```

```
QY 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTTACAGGAATGCA 1140
Db |||
1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTTACAGGAATGCA 1140
QY 1141 TTGTTTGTGCTCACTACAAACCGCACACAGCATATATCGATTGAGGGGACGG 1200
Db |||
1141 TTGTTTGTGCTCACTACAAACCGCACACAGCATATATCGATTGAGGGGACGG 1200
QY 1201 GCTCAGCTCAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
Db |||
1201 GCTCAGCTCAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
QY 1261 GGTCAAGTCTGTGTGTCACAGAACTTCGCCGTGCTGGAAAGTCCAGAGCGAGAAC 1320
Db |||
1261 GGTCAAGTCTGTGTGTCACAGAACTTCGCCGTGCTGGAAAGTCCAGAGCGAGAAC 1320
QY 1321 TTCGAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA 1380
Db |||
1321 TTCGAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA 1380
QY 1381 AACTCCGTCTATAGATAACCTCGCGGAGGAGTGTTCGAAATTCATATGGCTCCAAAGG 1440
Db |||
1381 AACTCCGTCTATAGATAACCTCGCGGAGGAGTGTTCGAAATTCATATGGCTCCAAAGG 1440
QY 1441 GAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Db |||
1441 GAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
Db |||
1501 CAGTCTCCGAGGCTGTGGCTTAA 1524

RESULT 2
AF093541
LOCUS Arachis hypogaea glycinin (Arah3) mRNA linear PLN 31-OCT-2001
DEFINITION Arachis hypogaea glycinin (Arah3) mRNA, partial cds.
ACCESSION AF093541
VERSION AF093541.1 GI:3703106
KEYWORDS
SOURCE
ORGANISM
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 1524)
REFERENCE
1 Rabjohn, P., Helm, E.M., Stanley, J.S., West, C.M., Sampson, H.A.,
AUTHORS Burks, A.W. and Bannion, G.A.
Molecular cloning and epitope analysis of the peanut allergen Ara h
3
J. Clin. Invest. 103 (4), 535-542 (1999)
JOURNAL
MEDLINE 99146968
PUBMED 10021462
REFERENCE
2 (bases 1 to 1524)
AUTHORS Rabjohn, P., Helm, E.M., Stanley, J.S., West, C.M., Sampson, H.A.,
Burks, A.W. and Bannion, G.A.
Direct Submission
TITLE Submitted (21-SEP-1998) Pediatrics, University of Arkansas for
JOURNAL Medical Sciences, 1120 Marshall Street, Little Rock, AR 72202, USA
FEATURES
Location/Qualifiers
source
1. .1524
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Florunner"
/db_xref="taxon:3818"
<1. .1524
/gene="Arah3"
<1. .1524
/gene="Arah3"
/feature="seed storage; peanut allergen"
/codon_start=1
/product="glycinin"
```


ORIGIN		/db_xref="taxon:3818"	
Query Match		99.1%; Score 1509.6; DB 6; Length 1524;	
Best Local Similarity		99.4%; Pred. No. 0;	
Matches 1515; Conservative		0; Mismatches 9; Indels 0; Gaps 0;	
Qy	1	CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCCAGCGCCTCAATGGCAGAGACCTGAC	60
Db	1	CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCCAGCGCCTCAATGGCAGAGACCTGAC	60
Qy	61	AATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACCAACGAGAGTTC	120
Db	61	AATCCCATTTGAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACCAACGAGAGTTC	120
Qy	121	GAATGGCGGGCTCGCCCTCTCTCGCTTAGTCTCGCGGCAACGCCCTTCGTAGGCTT	180
Db	121	GAATGGCGGGCTCGCCCTCTCTCGCTTAGTCTCGCGGCAACGCCCTTCGTAGGCTT	180
Qy	181	TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGATACATTTGGGTTG	240
Db	181	TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAAGGGATACATTTGGGTTG	240
Qy	241	ATATTCCCTGGTTGTTCTTAGACACTATGAAAGAGCCTCACACAAGGTTCGTGATCTCAG	300
Db	241	ATATTCCCTGGTTGTTCTTAGACACTATGAAAGAGCCTCACACAAGGTTCGTGATCTCAG	300
Qy	301	TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT	360
Db	301	TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGT	360
Qy	361	CACCAAGAGTGCACCGTTTCGATGAGGGTGATCTCAITGCACTTCCCAACCGGTGTGCT	420
Db	361	CACCAAGAGTGCACCGTTTCGATGAGGGTGATCTCAITGCACTTCCCAACCGGTGTGCT	420
Qy	421	TTCTGGCTCTAACAAGACACGACACTGATGTTGCTGTTTCTCTTACTGACACCAAC	480
Db	421	TTCTGGCTCTAACAAGACACGACACTGATGTTGCTGTTTCTCTTACTGACACCAAC	480
Qy	481	AACAACGACACACAGCTTGATAGTTTCCCAAGAGATTCATTTGCTGGGAACGGAG	540
Db	481	AACAACGACACACAGCTTGATAGTTTCCCAAGAGATTCATTTGCTGGGAACGGAG	540
Qy	541	CAAGAGTTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT	600
Db	541	CAAGAGTTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT	600
Qy	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660
Db	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGTGAAATTTAGCCCTCGAGGA	660
Qy	661	CAGCAGCGCGCAGAGAACGAGCAGCAAGAGAGAGAAACCAAGGTGGAACATCTTC	720
Db	661	CAGCAGCGCGCAGAGAACGAGCAGCAAGAGAGAGAAACCAAGGTGGAACATCTTC	720
Qy	721	AGCGGCTTACAGCCGGAGTTCTTCGGAACAGCCCTCCAGGTTACACACAGACAGATAGTG	780
Db	721	AGCGGCTTACAGCCGGAGTTCTTCGGAACAGCCCTCCAGGTTACACACAGACAGATAGTG	780
Qy	781	CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGA	840
Db	781	CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGA	840
Qy	841	GGCCTCAGATCTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGGAATACGAT	900
Db	841	GGCCTCAGATCTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAGGGAATACGAT	900
Qy	901	GAAGATGAATATGAATACGATGAAGAGGATAGAAGGGCTGGCAGGGGAAGCAGAGCAGG	960
Db	901	GAAGATGAATATGAATACGATGAAGAGGATAGAAGGGCTGGCAGGGGAAGCAGAGCAGG	960
Qy	961	GGGAATGGTATTGAAGAGAGGATCTCGACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC	1020
Db	961	GGGAATGGTATTGAAGAGAGGATCTCGACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC	1020
Qy	1021	AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCCAACGATCTCAAC	1080
Db	1021	AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCCAACGATCTCAAC	1080
Qy	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA	1140
Db	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA	1140
Qy	1141	TTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACCG	1200
Db	1141	TTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACCG	1200
Qy	1201	GCTCAGCTGCAAGTCTGTGACAGCAACGCAACAGAGTGTACGACAGAGAGCTTCAAGAG	1260
Db	1201	GCTCAGCTGCAAGTCTGTGACAGCAACGCAACAGAGTGTACGACAGAGAGCTTCAAGAG	1260
Qy	1261	GGTCAGCTGCTTGTGTGTCACAGAACTTTCGCGCTGCTGGAAAGTCCCGAGCGAGAAC	1320
Db	1261	GGTCAGCTGCTTGTGTGTCACAGAACTTTCGCGCTGCTGGAAAGTCCCGAGCGAGAAC	1320
Qy	1321	TTTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGTGAA	1380
Db	1321	TTTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGTGAA	1380
Qy	1381	AACTCGCTCATAGATAAAGTCCGCGGAGAGTGGTTGCAAAATTCATATGGCTTCCAAAGG	1440
Db	1381	AACTCGCTCATAGATAAAGTCCGCGGAGAGTGGTTGCAAAATTCATATGGCTTCCAAAGG	1440
Qy	1441	GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTTCTTGGTCCACCGTCTCAG	1500
Db	1441	GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTTCTTGGTCCACCGTCTCAG	1500
Qy	1501	CAGTCTCCGAGGCTGTGGCTTAA	1524
Db	1501	CAGTCTCCGAGGCTGTGGCTTAA	1524
RESULT 4			
AX148741		1853 bp	DNA linear PAT 08-JUN-2001
LOCUS		Sequence 4 from Patent WO0136621.	
DEFINITION		AX148741	
ACCESSION		AX148741.1	GI:14347295
VERSION			
KEYWORDS		Arachis hypogaea (peanut)	
SOURCE		Arachis hypogaea	
ORGANISM		Arachis hypogaea	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;	
		Aeschynomeneae; Arachis.	
REFERENCE		1	
AUTHORS		Dodo, H.W., Arntzen, C.J., Konan, K.N. and Viquez, O.M.	
TITLE		Down-regulation and silencing of allergen genes in transgenic	
		peanut seeds	
JOURNAL		Patent: WO 0136621-A 4 25-MAY-2001;	
		Alabama A & M University (US)	
FEATURES		Location/Qualifiers	
source		1..1853	
		/organism="Arachis hypogaea"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3818"	
ORIGIN			
Query Match		91.3%; Score 1391.2; DB 6; Length 1853;	
Best Local Similarity		95.4%; Pred. No. 0;	
Matches 1454; Conservative		0; Mismatches 68; Indels 2; Gaps 2;	
Qy	2	GGCAGCAACCGGAGAGAACGGTGCAGTTCAGCGCCTCAATGCGCAGAGACCTGACA	61
Db	71	GGCAGCAACCGGAGAGAACGGTGCAGTTCAGCGCCTCAATGCGCAGAGACCTGACA	130

QY 62 ATGCATTGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACGAGGTTGC 121
Db 131 ACCGATTTGAATCGAGGCGGTTACATTGAGACTTGGAAACCCCAACACGAGGTTGC 190
QY 122 AATCGCGCGGCGTCCGCTCTCTCGCTTAGTCTCTCGCGCAACGCGCTTCTGAGGCTTT 181
Db 191 AATCGCGCGGCGTCCGCTCTCTCGCTTAGTCTCTCGCGCAACGCGCTTCTGAGGCTTT 250
QY 182 TCTACTCAATGCTCCGAGGAGATCTTCATCCAGCAAGGAAGGAGATCTTTGGGTTGA 241
Db 251 TCTACTCAATGCTCCGAGGAGATCTTCATCCAGCAAGGAAGGAGATCTTTGGGTTGA 310
QY 242 TATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCACACAGAGGTCGTCGATCTCACT 301
Db 311 TATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCACACAGAGGTCGTCGATCTCACT 370
QY 302 CCAAAGACCAACGAGAGCTTCACAGGAGAGAACCAAGAGCCAAACAGCAACGAGATAGTC 361
Db 371 CCAAAGACCAACGAGAGCTTTGCAAGAGAGAGAACCAAGAGCCAAACAGCAACGATAGTC 430
QY 362 ACCAAGAGGTCACCGTTTGATGAGGAGTATCTCATTTGCGATTTCCACCGGTTGCTTT 421
Db 431 ACCAAGAGGTCACCGTTTCAATGAGGAGTATCTCATTTGCGATTTCCACCGGTTGCTTT 490
QY 422 TCTGGCTCTAACAGCACGACGACTGATGTTGTTGTTGTTGTTCTTCTTACTGACACCAACA 481
Db 491 TCTGGCTGTAAAGACACGACGACTGATGTTGTTGTTGTTGTTCTTCTTACTGACACCAACA 550
QY 482 ACAACGACCAACGAGCTTGATCAGTTCCCGAGGAGATTCATTTTGGCTGGGAACGCGAGC 541
Db 551 ACAACGACCAACGAGCTTGATCAGTTCCCGAGGAGATTCATTTTGGCTGGGAACGCGAGC 610
QY 542 AAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATA 601
Db 611 AAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATA 670
QY 602 GCCCATACGCGCAAGTCAGCTAGACAGAGAGCGTGATTTAGCCCTCGAGGAC 661
Db 671 GCCCATACGCGCGATAGTCGGCTAGACAGAGAGCGTGATTTAGCCCTCGAGGAC 730
QY 662 AGCACGCGCAGAGAACGAGCAGACAAAGAGAGAGAAACGAGGTGGAAACATCTTCA 721
Db 731 AGCACGCGCAGAGAACGAGCAGACAAAGAGAGAGAAACGAGGTGGAAACATCTTCA 790
QY 722 CGCGCTTCAAGCGGAGTTCTTGAAACAGCTTCCAGGTGACGACAGACAGATAGTC 781
Db 791 CGCGCTTCAAGCGGAGTTCTTGAAACAGCTTCCAGGTGACGACAGACAGATAGTC 850
QY 782 ABAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGCCATTTGACAGTGAAGGAG 841
Db 851 ABAATCTGTGGGCGAGAACGAGAGTGAAGAGAGGAGCCATTTGACAGTGAAGGAG 910
QY 842 GCCTCAGAACTTTGAGCGCCAGATAGAGAGAGAGAGAGAGAGAGAGAGAGATACGATG 901
Db 911 GCCTCAGAACTTTGAGCGCCAGATAGAGAGAGAGAGAGAGAGAGAGAGATACGATG 970
QY 902 AAGATGAATATGAATATGATGAAGAGATGAAGCGCTGCGAGGGAAGCAGAGCGAGG 961
Db 971 AAGATGAATATGAATATGATGAAGAGATGAAGCGCTGCGAGGGAAGCAGAGCGGCG 1030
QY 962 GGAATGCTATTGAAGAGAGAGATCTGACCGCAAGTGTCTTAAAGAGACATTTGTAAGA 1021
Db 1031 GGAATGCTATTGAAGAGAGAGATCTGACCGCAAGTGTCTTAAAGAGACATTTGTAAGA 1090
QY 1022 GATCCCTGACATCTAACACCTCA-AGCTGGTTTCACTCAAAACTGCGCAACGATCTCAAC 1080
Db 1091 GATCCCTGACATCTAACATCTCAGCGCTGTTTCACTCAAAACTGCGC-ACGATCTCAAC 1149
QY 1081 CTTCTAATATCTAGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTTCTACAGGAATGCA 1140
Db 1150 CTTCTAATCTTCTAGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTTCTACAGGAATGCA 1209
QY 1141 TTGTTTTCGCTCACTACAAACCAACGACACAGCATCATATATCATTCGATTGAGGGGACGG 1200

Db 1210 TTGTTTTCGCTCACTACAAACCAACGACACAGCATCATATATGATTTGAGGGGACGG 1269
QY 1201 GCTCAGTGCAGTTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1270 GCTCAGTGCAGTTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1329
QY 1261 GGTCAAGTTCGTTCGTGGCCACAGAACTTCGCGTTCGTGGAAAGTCCGAGAGGAGAAC 1320
Db 1330 GGTCAAGTTCGTTCGTGGCCACAGAACTTCGCGTTCGTGGAAAGTCCGAGAGGAGAAC 1389
QY 1321 TTGCAATACGTGTCATTTCAAGACAGACTCAAGGCCACGATAGCCAACTTCGCGGTGAA 1380
Db 1390 TTGCAATACGTGTCATTTCAAGACAGACTCAAGGCCACGATAGCCAACTTCGCGGTGAA 1449
QY 1381 AACTCGCTCATAGATAACCTGCGGAGGAGGTGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 1450 AACTCGCTCATAGATAACCTGCGGAGGAGGTGTTGCAAAATTCATATGCGCTCCCAAGG 1509
QY 1441 GAGCAGCAGGCGAGCTTAAGAACAAACCCCTTCAGTTCCTGTTCCACCGCTCTCAG 1500
Db 1510 GAGCAGCAGGCGAGCTTAAGAACAAACCCCTTCAGTTCCTGTTCCACCTTTTCAG 1569
QY 1501 CAGTCTCCGAGGCGCTGTGGCTTAA 1524
Db 1570 CAGTCTCCGAGGCGCTGTGGCTTAA 1593

RESULT 5
AF086821
LOCUS 1853 bp mRNA linear PLN 29-SEP-1999
DEFINITION Arachis hypogaea glycinin (AraH4) mRNA, complete cds.
ACCESSION AF086821
VERSION AF086821.1 GI:5712198
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 1853)
AUTHORS Kleber-Janke,T., Cramer,R., Appenzeller,U., Schlaak,M. and Becker,W.M.
TITLE Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology
JOURNAL Int. Arch. Allergy Immunol. 119 (4), 265-274 (1999)
MEDLINE 99406463
PUBMED 10474031
REFERENCE 2 (bases 1 to 1853)
AUTHORS Kleber-Janke,T.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-1998) Biochemische und Molekulare Allergologie, Forschungszentrum Borstel, Parkallee 22, Borstel 23845, Germany
FEATURES
source 1. .1853
/organism="Arachis hypogaea"
/mol_type="mRNA"
/strain="Virginia"
/db_xref="taxon:3818"
/dev_stage="seed"
1. .1853
/genes="AraH4"
1. .1593
/notes="seed storage protein; allergen"
/codon_start=1
/product="glycinin"
/protein_id="AAD47382.1"
/db_xref="GI:5712198"
/translation="MAKLELSPFCFLVLGASSISFROOPEBNAQCFORLNAORPDN
RISEGGYLETWNPNNQEPCEAGVALSLRLRNALRRPFSNAPQEIFIQCGRGVFG
LIFPGCPSTYEPAQQGRYQSORPPRLQEEDSQSQDQSHQKVRHFNEDGLIAVPT

GVAFWLYNDHDTDVAVLSVLTNNNDNQLDQPPRRFNLNAGNHEQEFRLRYQOOSRQR
RSLPSPYSPHSRPREERFRPRGRAGQEEDEEGNIFSGFTPEFLEQAFQ
VDRQIIVONLWGNSEBEGAIIVTGGRLILSPDTRGADBEEDYDFQYEVHEODG
RCGRSGRGGNGIEETICTACVKKNIGWNEPHIYDPQRWFTQNDHDLNLLILWLGL
SAGYLNLYRNALFVPHYNTNAHSIIIVALGRHAHVQVDSNGRVYDDEELQEGHVLVVP
QNFPAVAKGQSENFVAFKTDSPRSIANFAGNSFIDNLPEEVVANSYGLPREQARQ
LKNPNFKFVPFPFOOSPRAVA"

ORIGIN

Query Match	91.4%;	Score	1391.2;	DB	8;	Length	1853;
Best Local Similarity	95.4%;	Pred. No.	0;				
Matches	1454;	Conservative	0;	Mismatches	68;	Indels	2;
Gaps	2;						
QY	2	GGCAGCAACCGGAGGAGACGCTGCGCAGTTCAGGCGCTCAATGGCGCAGACCTTGACA	61				
DB	71	GGCAGAGCCGGAGGAGATGCGTCCAGTTCAGGCGCTCAATGGCGCAGACCTTGACA	130				
QY	62	ATCGCAATGAATCAGAGGGCGGTTACATTCAGACTTTGGAAACCCCAACAAACAGGAGTTGC	121				
DB	131	ACGCGATTGAATCGGAGGGCGGTTACATTCAGACTTTGGAAACCCCAACAAACAGGAGTTGC	190				
QY	122	AATGGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCGCTTCGTAGGCGTT	181				
DB	191	AATGGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCGCTTCGTAGGCGTT	250				
QY	182	TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTTGA	241				
DB	251	TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTTGA	310				
QY	242	TATTCCCTGTTGCTCTAGACACTATGAGAGCCTCACACAAAGTGCTGTCAGT	301				
DB	311	TATTCCCTGTTGCTCTAGACACTATGAGAGCCTCACACAAAGTGCTGTCAGT	370				
QY	302	CCCAAGACCCACAGACGCTCCAGGAGAACCAAGCCCAACAGCAGAGATAGTC	361				
DB	371	CCCAAGACCCACAGACGTTTGCAGAGAGAACCAAGCCCAACAGCAGAGATAGTC	430				
QY	362	ACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCGATTCGCCCGGTTGCTT	421				
DB	431	ACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCGATTCGCCCGGTTGCTT	490				
QY	422	TCTGGCTCTACAAACGACGACACTGATGTTGCTGTTTCTCTTACAGACCAACA	481				
DB	491	TCTGGCTCTACAAACGACGACACTGATGTTGCTGTTTCTCTTACAGACCAACA	550				
QY	482	ACAACGACACACGAGTGTATCAGTTCGCCAGGAGATTCATTTGGCTGGGAACCGGAC	541				
DB	551	ACAACGACACACGAGTGTATCAGTTCGCCAGGAGATTCATTTGGCTGGGAACCGGAC	610				
QY	542	AAGAGTTCTTAAGGTACCGACCAAGACGACAGCAAAAGCAGCAAGAGCTTACCATATA	601				
DB	611	AAGAGTTCTTAAGGTACCGACCAAGACGACAGCAAAAGCAGCAAGAGCTTACCATATA	670				
QY	602	GCCCATACAGCCGCAAGTACGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGAC	661				
DB	671	GCCCATACAGCCGCGATAGTCGCGCTAGACAGAGAGCGTGAATTTGCGCTCGAGGAC	730				
QY	662	AGCACAGCCGACAGACGAGCAGGACAGAAAGAAACGAGGTGGAAACATCTTCA	721				
DB	731	AGCACAGCCGACAGACGAGCAGGACAGAAAGAAACGAGGTGGAAACATCTTCA	790				
QY	722	GCGGCTTCACGCGGAGTTCTCTGGAACAAGCCCTTCAGGTTGACGACAGACATAGTGC	781				
DB	791	GCGGCTTCACGCGGAGTTCTCTGGAACAAGCCCTTCAGGTTGACGACAGACATAGTGC	850				
QY	782	AAACCTAAGAGCGGACAGCGAGGTGAAGAAAGAGGAGCCATTGTGACAGTGGGGAG	841				
DB	851	AAAATCTGTGGGCGAAGACGAGAGTGAAGAGAGGGAGCCATTGTGACGCTGAGGGAG	910				
QY	842	GCCTCAGATCTTGAGCCCGATAGAAAGAGAGCTGCCGACGAGAGAGGATAGATG	901				
DB	911	GCCTCAGATCTTGAGCCCGATAGAAAGAGAGCTGCCGACGAGAGAGGATAGATG	970				

QY	902	AAGATGAATATGAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAACGAGGACGG	961				
DB	971	AAGATCAATATGAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAACGAGGACGG	1030				
QY	962	GGAAATGGTATTGAAGAGAGCATCTGCACCGCAAGTGTCTAAAAAGAAACATTGGTAGAAACA	1021				
DB	1031	GGAAATGGTATTGAAGAGAGCATCTGCACCGCATGTGTAAAAAGAAACATTGGTAGAAACA	1090				
QY	1022	GATCCCTCGACATCTACAAACCTCTCA-AGCTGTGTTCACTCAAAACCTGCCAAACATCTCAAC	1080				
DB	1091	GATCCCTCGACATCTACAAACCTCTCAAGCTGTGTTCACTCAAAACCTGCC-ACGATCTCAAC	1149				
QY	1081	CTTCTAATACTTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA	1140				
DB	1150	CTTCTAATCTTAGTGGCTTGGACTTAGTGTGTAATATGGAATCTCTACAGGAATGCA	1209				
QY	1141	TTGTTTGTGCTCTACATAACCAACCAACGCAACAGCATCATATATCGATTGAGGGACGG	1200				
DB	1210	TTGTTTGTGCTCTACATAACCAACCAACGCAACAGCATCATATATGAGGGACGG	1269				
QY	1201	GCTCAGTGCAGAGTGTGGACAGACGCAACAGAGTGTACGAGGAGCTTCAAGAG	1260				
DB	1270	GCTCAGTGCAGAGTGTGGACAGACGCAACAGAGTGTACGAGGAGCTTCAAGAG	1329				
QY	1261	GCTCAGTGTGTTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCGAGAGCGAGAAC	1320				
DB	1330	GCTCAGTGTGTTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCGAGAGCGAGAAC	1389				
QY	1321	TTGCAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCGCGGTGAA	1380				
DB	1390	TTGCAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTTTCCGCGGTGAA	1449				
QY	1381	AACTCCGTCATAGATAACCTCGCGGAGGAGTGTGTTGCAAAATTCATATGGCTTCCAAAGG	1440				
DB	1450	AACTCCGTCATAGATAACCTCGCGGAGGAGTGTGTTGCAAAATTCATATGGCTTCCCAAGG	1509				
QY	1441	GAGCAGGCAAGCGCTTAAGAACCAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG	1500				
DB	1510	GAGCAGGCAAGCGCTTAAGAACCAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG	1569				
QY	1501	CAGTCTCCGAGGCTGTGCTTAA	1524				
DB	1570	CAGTCTCCGAGGCTGTGCTTAA	1593				
RESULT	6						
AF125192							
LOCUS							
DEFINITION	Arachis hypogaea Gly1 (Gly1) mRNA, linear	1590 bp	mRNA	linear	PLN 22-AUG-2000		
ACCESSION	AF125192						
VERSION	AF125192.2						
KEYWORDS	GI:9864776						
SOURCE	Arachis hypogaea (peanut)						
ORGANISM	Arachis hypogaea						
TITLE	Arachis hypogaea						
REFERENCE	Arachis hypogaea						
AUTHORS	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;						
JOURNAL	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;						
AUTHORS	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;						
JOURNAL	Aeschynomeneae; Arachis.						
REFERENCE	1 (bases 1 to 1590)						
AUTHORS	Jain,A.K. and Basha,S.M.						
JOURNAL	Molecular cloning of a glycinein-type peanut seed storage protein						
REFERENCE	2 (bases 1 to 1590)						
AUTHORS	Jain,A.K. and Basha,S.M.						
JOURNAL	Direct Submission						
REFERENCE	3 (bases 1 to 1590)						
AUTHORS	Jain,A.K. and Basha,S.M.						
JOURNAL	Submitted (02-FEB-1999) Plant Biotechnology, Florida A&M						
REFERENCE	USA						
AUTHORS	University, 301 South Perry Paige Building, Tallahassee, FL 32307,						
JOURNAL	Submitted (22-AUG-2000) Plant Biotechnology, Florida A&M						
REFERENCE	USA						
AUTHORS	University, 301 South Perry Paige Building, Tallahassee, FL 32307,						
JOURNAL	Submitted (22-AUG-2000) Plant Biotechnology, Florida A&M						
REFERENCE	USA						

[illegible]

Db 1392 TGGCTGGAAGTCCCGAGGAGAGAACTTTGAAATAGCTGGGCATTTCAAGACAGACTCAAGGC 1451

Qy 1355 CCAGCATAGCCAACTCTCCCGGTGAAAACTCCGTCATAGATAAACCCTCCGAGGAGGTGG 1414

Db 1452 CCAGCATAGCCAACTAGCCCGTGAAAACCTCTTCATAGATAAATTGCGCGAGGAGGTGG 1511

Qy 1415 TTGCAAAATTCATATGGGCTCCAAAGGAGGAGGCAAGGCAGCTTTAAGAACAAACACCCCT 1474

Db 1512 TTGCAAAATTCATATGGGCTCTCCAGGAGGAGGCAAGGCAGCTTTAAGAACAAACACCCCT 1571

Qy 1475 TCAAGTTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524

Db 1572 TCAAGTTCTTCGTTCCACCGTCTGAACAGTCTCTGAGGGCTGTGGCTTAA 1621

RESULT 10

AY618460 1533 bp mRNA linear PLN 07-JUN-2004

LOCUS Arachis hypogaea glycinin mRNA, partial cds.

DEFINITION Arachis hypogaea glycinin mRNA, partial cds.

ACCESSION AY618460

VERSION Arachis hypogaea

KEYWORDS Arachis hypogaea (peanut)

SOURCE Arachis hypogaea

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae; Arachis.

REFERENCE 1 (bases 1 to 1533)

AUTHORS Kang,I.-H. and Gallo-Weagher,M.

TITLE Cloning and characterization of a novel ara h 3, a major peanut (Arachis hypogaea L.) allergen gene

JOURNAL Unpublished

AUTHORS 2 (bases 1 to 1533)

TITLE Kang,I.-H. and Gallo-Weagher,M.

JOURNAL Direct Submission

TITLE Submitted (05-MAY-2004) Agronomy, University of Florida, Gainesville, FL 32611, USA

FEATURES

source 1..1533

1 /organism="Arachis hypogaea"

2 /mol_type="mRNA"

3 /db_xref="taxon:3818"

4 <1..1533

5 /notes="seed storage protein; Ara h 3-im; peanut allergen"

6 /codon_start=1

7 /product="glycinin"

8 /protein_id="AAT39430.1"

9 /db_xref="GI:47933675"

10 translations="KIALSLICFVLVGASVTFROGGENEQFORLNAORPDNRI

11 ESQGIETWPNNOIEPQCAVGLSRVLRNALRRFYSNAPLELYVOOGSGVGLI

12 PFGCPSTYEPAGRGYQSKSRRFQVQDDPSQQQDSHQKVRHFRDGLLIAPVT

13 GVAFMTNDDDVVTLLSDTSIHNQDQFPRFYLAVGNQSFQRYOQQQSGRPH

14 YRQISPRVDDEGENEENITFSGFAQFQHFQVDRQTVENLRNRYKQEQGSRPH

15 KGLRILSPDEDESSPPSRREEFDDRSRQQRKYDNRNRGKNGYEETTCAS

16 VKNLGRSNPDITYNPAQSLRSVNELDPLILGWLGAOHGTIYRAMEVPVHTLNA

17 HTIIVALNGRAHVQVDSNGRVYDEIQSHVLVWPQFAVAAKAQSENVEYLAFT

18 DSRPSIANLAGENSIIDNLEEVVANSYRLPREQARQLKNNPFKFPVFPFDQSMRE

19 VA"

ORIGIN

Query Match 56.0%; Score 854.2; DB 8; Length 1533;

Best Local Similarity 76.1%; Pred No. 1.8e-212;

Matches 117; Conservative 0; Mismatches 288; Indels 102; Gaps 6;

Qy 2 GGCAGCAACCGGAGGAGAACCGGTGCCAGTTCGACGCGCTCAATGCGCAGAGACTGACA 61

Db 65 GGCAGGGGGGAGGAGGAATGAGTGCCAGTTCAGCGCTCAATGCGCCAGAGACCCGACA 124

Qy 62 ATCGCATTGAATCAGAGGGCGGTTACATTGAGACTTGGAACCCCAACAACAGAGGATTGC 121

Db 125 ACCGCATTGAGTCTGAAGGGCGGTTACATTGAGACTTGGAACCCCAACAACAGGAGTTCC 184

Db 1187 TTGTGGCATTGAACGGACGGGCACATGTGCAAGTGGTGGACGCAACCGGTAAACAGAGTGT 1246
Qy 1241 ACACGAGGAGCTTCAAGAGGGTCACTGCTGCTTGTGGTGGCCACAGAACTTCGCCGTGCGCTG 1300
Db 1247 ACCACGAGGAGCTTCAAGAGGGTCACTGCTGCTTGTGGTGGCCACAGAACTTCGCCGTGCGAG 1306
Qy 1301 GAAAGTCCACAGAGGAGAACTTCGAATACCTGGCATTTCAAGACAGACTCAAGGCCACAGCA 1360
Db 1307 CAAAGGCCACAGAGCGAGAACTACGAATACCTGGCGTTCAAGACAGACTCAAGGCCACAGCA 1366
Qy 1361 TAGCCAACTCCCGGTGAAATCCCTCATAGATACCTGCCGAGGAGGTGGTTCGAA 1420
Db 1367 TAGCCAACTCCCGGTGAAATCCCTCATAGATACCTGCCGAGGAGGTGGTTCGAA 1426
Qy 1421 ATTATATGCTCCCAAGGGGAGCAGGCAAGGAGCTTAAGAAACAACCCCTTCAAGT 1480
Db 1427 ATTCTTACCGCTCCCAAGGGGAGCAGGCAAGGCACTTAAGAAACAACCCCTTCAAGT 1486
Qy 1481 TCTTGGTTCACCGT---CTCAGCAGTCTCCGAGGGGTGTGGCTTAA 1524
Db 1487 TCTTGGTTCACCGTTCGATCATCAGTCTATGAGGGAGGTGGCTTAA 1533

RESULT 11
AF510854
LOCUS AF510854 3825 bp DNA linear PLN 03-JUN-2002
DEFINITION Arachis hypogaea allergen Arah3/Arah4 gene, complete cds.
ACCESSION AF510854
VERSION AF510854.1 GI:21314464
KEYWORDS
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 3825)
Viquez,O.M., Konan,K.N. and Dodo,H.W.
Genomic characterization of the major peanut allergen genes, Arah3
and/or Arah4
Unpublished
REFERENCE 2 (bases 1 to 3825)
Viquez,O.M., Konan,K.N. and Dodo,H.W.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2002) Food and Animal Sciences, Alabama A&M
University, Meridian Street, 4900, Normal, AL 35762, USA
FEATURES
source
1. .3825
/organism="Arachis hypogaea"
/mol_type="genomic DNA"
/db_xref="taxon:3818"
TATA_signal
mRNA 597..603
join(<661..949,1289..1581,1669..2313,2482..3825)
/product="allergen Arah3/Arah4"
CDS join(661..949,1289..1581,1669..2313,2482..2871)
/codon_start=1
/product="allergen Arah3/Arah4"
/protein_id="AAM46958.1"
/db_xref="GI:21314465"
/translations="MGKLLALSVCFLVLGASSISFRQOPPENACQFORLNAQRPDN
RISEGGYIEITWPNNOEFECAGVALSLRLVLRNALRPFYSNAPQEIFIQQRGYFG
LIPFGCPSTVEEPAQGRHQSRPPRRFQGDQSQDQSHQKVRHFDGDLIAVPT
GVAFWMNDHTDVAVSLTDNNNDQLDQPPRFLNAGNHEQFLRYQQQSRRLS
PYSYSPQTPQKDEEFEPGRQGRERAGOEENGNLFSGFTPEFLAQFOVD
RQLQLRGENESEDGALIVRGGRLILSPDKRQQQYERDEEYDEYDEE
ERQDRLRRGSGSGNGLEETICTASFKKNIGNRSPDIYNPQASLKTANELQNL
LILRWLGLSALYRNALFVPHYNTNAHSIIYALRGRAHVQVDSNGRVPDEELQ
EGHVLVVPQNFVAGKSQSENFEYVAFKTDSPRSIANLAGENSFIDNLPEEVVANSYG
LPREAROLKNNNPFFKFFVPPSPQSURAVA"
<661..949
/number=1
1289..1581

exon /number=2 1669..2313
exon /number=3 2482..3825
3'UTR /number=4 2872..3825
ORIGIN
Query Match 54.8%; Score 834.6; DB 8; Length 3825;
Best Local Similarity 77.1%; Pred. No. 2.6e-207;
Matches 1233; Conservative 0; Mismatches 69; Indels 297; Gaps 6;
Qy 214 CAGCAAGGAGGGGATACCTTTGGGTGTGATATTTCCCTCGTGTTCCTAGACACTATGAAGAG 273
Db 1282 CATAAGGAGGGGATACCTTTGGGTGTGATATTTCCCTCGTGTTCCTAGACACTATGAAGAG 1341
Qy 274 CTTCAACACCAAGGTCGTGATCTCAGTCCCAAGACCAACAGAGCTCTCAAGGAGAA 333
Db 1342 CTTGCAACAAGGAGCGCCGACATCATGTCCCAAGACCAACAGAGCTTTCAAGGACAA 1401
Qy 334 GACCAAGCCAAACAGCAACAGAGATAGTCAACCAAGGTCACCGTTTCGATGAGGCTGAT 393
Db 1402 GACCAAGCCAAACAGCAACAGAGATAGTCAACCAAGGTCACCGTTTCGATGAGGCTGAT 1461
Qy 394 CTCATTGCAAGTCCCAACCGGTGTGCTTCTGGCTCTACAACGACCAACGACTGATGTT 453
Db 1462 CTCATTGCAAGTCCCAACCGGTGTGCTTCTGGATGTACAACGACCATGACTGATGTT 1521
Qy 454 GTTGTGTTTCTTCTTACTGACCAACCAACAGCAACCAAGCTTGATCAGTTTCCCCAG- 512
Db 1522 GTTGTGTTTCTTCTTACTGACCAACCAACAGCAACCAAGCTTGATCAGTTTCCCCAGG 1581
Qy 513 ----- 512
Db 1582 GTATATATATGTTTCAATCTCTATTATTTTGGCATGTGCATATATATCATGAAGGA 1641
Qy 513 -----GAGATTCAATTTGGCTGGGAAACGAGGAGCAAGAG 546
Db 1642 ATGGTGACAATTAATAATACATAACAGAGATTCAATTTGGCTGGGAAACGAGGAGAG 1701
Qy 547 TTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAAGAGCTTACCATATAGCCCA 606
Db 1702 TTCTTAAGATACCAGCA-----ACAAGCAGCAAGAGCTTACCATATAGCCCA 1752
Qy 607 TACAGCCGCCAAAGTCAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGACACAC 666
Db 1753 TACAGCCGCCAAACTCAGCCTTAAACAAGAGACCGTGAATTTAGCCCTCGAGGACACAC 1812
Qy 667 AGCCGACAGAACGAGCAGGACAGAAAGAAACGAAAGGTGGAAAATCTTTCAGCGGC 726
Db 1813 GGCCGACAGAACGAGCAGGACAGAAACGAAAGGTGGAAAATCTTTCAGCGGC 1872
Qy 727 TTCAAGCGCGAGTCTCTGGAAACAAGCTTCCAGGTTTGACGACAGACAGATAGTGCAAAAC 786
Db 1873 TTCAAGCGCGAGTCTCTGGCAAGCTTCCAGGTTTGACGACAGACAGATATTTGCAAAAC 1932
Qy 787 CTAAGAGGCGCAGACCCGAGAGTGAAGAGAGGAGGCCATTTGTGACAGTGAAGGAGGCGCTC 846
Db 1933 CTAAGAGGCGCAGACGAGAGTGCACAGGAGGCCATTTGTGACAGTGAAGGAGGCGCTC 1992
Qy 847 AGAATCTTGAAGCCAGATGAAGAGAG-----ACGTGCCCAAGAAAGAG 891
Db 1993 AGAATCTTGAAGCCAGATGAAGAGAGAGCGCAGCAGTATGAACGTCCTCCGACGAAAGAG 2052
Qy 892 GAATCAGATGAAGATGAATATCAATACGATGAAGAG-----CATAGAAGCGT 939
Db 2053 GAATCAGATGAAGATGAATATCAATATGAATGAAGAGAGGAGGCAACAAGATAGAGCGCT 2112
Qy 940 GGCAGGGGAGCAGAGGCGAGGGGGAATGGTATTTGAAGAGACGATCTGCAACCGCAAGTGCT 999
Db 2113 GGCAGGGGAGCAGAGGCGAGCGCAATGGCATTTGAGGAGACCATCTGCAACCGCAAGTTT 2172
Qy 1000 AAAAAAGAACATTTGGTAGAAAAACAGATCCCTCTGACATCTACAACTCAAGCTGGTTCACTC 1059

Db	739	GCGGCTTCTCAACAGATTCTCTCAATGGCTTCCAGTGAACGA---GGACATAGTGA	795
QY	782	AAAACTAAGAGCGGACGACGAGAGTGAAGAAGAGGAGCCATTGTCAAGTGAAGGAG	841
Db	796	GGAACTTAAGAGGAGAAAACGACGCGAGGAGCAAGGAGCCATTGTCAAGTAAAGGAG	855
QY	842	GCCTCAGAATCTTTAGCCAGATAGAAAAGAGAGCTGCGGACGAAAGAGGAATACGATG	901
Db	856	GACTTAGCATCTTAGTTTCCACAGAGTGGAGGAGAGCTACCAACAA-----	902
QY	902	AAGATGAATGAATACGATGAAGAGATAGAGGCGTGCAGGGGAGCAGAGGAGG	961
Db	903	-----	927
QY	962	GGATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAAAGAACATTGGTAGAAACA	1021
Db	928	ACAAAGGCAATTGAAGAAACCAATTGCACCTGTTAAGATGAACATTGGCAATCCA	987
QY	1022	GATCCCTGACATCTCAACACCTCAAGCTGGTTTCATCTCAAACTGCCAAGCATCTCAACC	1081
Db	988	CATCTGTGATATCTCAACACCTCAAGCTGGTAGCGTCAAGACTGTAAACGAACCTTGACC	1047
QY	1082	TTCTAATACCTTAGTGGCTTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCAT	1141
Db	1048	TCCAAATCCTCAATCGACTCGGACTCAGTGCAGATGGATGCCATTCTATCGGGATGCAA	1107
QY	1142	TGTTTGTGCTCACTACAAACCAACGACACAGCATCATATCGATTGAGGGGACGGG	1201
Db	1108	TGTTTGTCTCACTACAACTGAACGCAACAGCATGATATGATTCACGCGAGGAG	1167
QY	1202	CTCAGTGAAGTCTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGG	1261
Db	1168	CTCATGTCCAAAGTGGTGGACTGCAATGGCAATAGAGTGTTCGACGAGGAGCTTCAAGAGG	1227
QY	1262	GTCAAGTGTCTGTGGTGCACAGAACTTCGCGTGCCTGGAAGTCCGACGAGGAGCACT	1321
Db	1228	GTCAATCGCTGGTGGTGCACAAACTTCGCGTGCCTGGAAGTCAACAGGAGGACCT	1287
QY	1322	TCCAAATCGTGGCACTCAAGACAGACTCAAGGCCACAGCATAGCAACCTCGCGGTGAAA	1381
Db	1288	TCCTTGTAGTGGATTCAGACAAACTCAAGGCCAGCATATCAACCTTGTGGCAAA	1347
QY	1382	ACTCGGTGATAGATAACCTGCGGAGGAGTGGTTGCAATTCAATATGCTCCAAAGGG	1441
Db	1348	ATTCTACATAGTGGAACTTGCCGGAAGATGTGGTTGCAAAATCATATGCTACAAATAG	1407
QY	1442	AGCAGGCAAGGAGCTTGAACCAACACCCCTCAAGTCTTCGTTCCACC	1493
Db	1408	AGCAAGCAAGGCAACTCAAGAACAAATAACCCCTTCAAGTCTTCGTTCCACC	1459
RESULT 13			
AF487543			
LOCUS	AF487543	657 bp	mRNA linear PLN 02-APR-2004
DEFINITION	Arachis hypogaea trypsin inhibitor mRNA, partial cds.		
ACCESSION	AF487543		
VERSION	AF487543.1 GI:22135347		
KEYWORDS	Arachis hypogaea (peanut)		
SOURCE	Arachis hypogaea		
ORGANISM	Arachis hypogaea		
REFERENCE	Dodo,H.W., Viquez,O.M., Maleki,S.J. and Konan,K.N.		
AUTHORS	cDNA clone of a putative peanut (Arachis hypogaea L.) trypsin		
TITLE	inhibitor has homology with peanut allergens Ara h 3 and Ara h 4		
JOURNAL	J. Agric. Food Chem. 52 (5), 1404-1409 (2004)		
PUBMED	14995153		
REFERENCE	2 (bases 1 to 657)		
AUTHORS	Dodo,H.W. and Viquez,O.M.		
TITLE	Direct Submission		

JOURNAL	Submitted (25-FEB-2002) Food and Animal Sciences, Alabama A&M University, Meridian Street, 4900, Normal, AL 35762, USA			
FEATURES	Location/Qualifiers			
source	1..657			
	/organism="Arachis hypogaea"			
	/mol_type="mRNA"			
	/db_xref="taxon:3818"			
	<1..>657			
CDS	/codon_start=1			
	/product="trypsin inhibitor"			
	/protein_id="AAM93157.1"			
	/db_xref="GI:22135348"			
	/translation="NYLHMLLALSVCFLVLGASSISFRQPEENACQFOLNLRQRP			
	DNRIESEGYIETWNNQNEFECAGVALSLVLRNLRPFYSNAPQEIFTQQRGY			
	FGLIFPGCPSTVEEPQQRHRSORPPRRFOGQSOQOQSDSHQKVRHFRDGLIAV			
	PTGVAFWMYNDHDTDVAVSLTDTNNNDNQDQPPRRFNLAGNHEQEFLRYQQQSRRR			
	S"			
ORIGIN				
Query Match	34.1%;	Score 519.2;	DB 8;	Length 657;
Best Local Similarity	93.4%;	Pred. No. 1e-124;		
Matches 542;	Conservative 0;	Mismatches 38;	Indels 0;	Gaps 0;
QY	2	GGCAGCAACGGAGAGAAACGGTGCAGTTCAGGGCCTCAATGGCGAGACCTGACA	61	
Db	77	GGCAGCAGCGGAGGAAATGCGTCCAGTTCAGCGCCTCAATGCGAAAGGCGCTGACA	136	
QY	62	ATCGCATTTGAATCAGAGGCGGTTCATTGACATTTGGAACCCCAACACAGGAGTTG	121	
Db	137	ACCGCATTTGAATCGGAGGCGGTTCATTGACATTTGGAACCCCAACACAGGAGTTG	196	
QY	122	AATGCGCGGGCTCGCCCTCTCTCGTTAGTCTCTCGCGCAACGCCCTTCGTAGGCTT	181	
Db	197	AATGCGCGGGCTCGCCCTCTCTCGTTAGTCTCTCGCGCAACGCCCTTCGTAGGCTT	256	
QY	182	TCTACTCAATGCTCCGAGGAGATCTTCATCCAGAAAGGGAATCTTTGGTTGA	241	
Db	257	TCTACTCAATGCTCCGAGGAGATCTTCATCCAGAAAGGGAATCTTTGGTTGA	316	
QY	242	TATTCCTCGTTGCTTAGACACTATGAAGAGCCTCACACAAAGTCTGTCTCAGT	301	
Db	317	TATTCCTCGTTGCTTAGACACTATGAAGAGCCTGCACAAAGGAGCCGACATCAGT	376	
QY	302	CCCAAGACCAACCAAGAGCGTCTCCAAGGAGAGACCAAGGCCCAACAGCATAGTC	361	
Db	377	CCCAAGACCAACCAAGAGCGTCTCCAAGGAGAGACCAAGGCCCAACAGCATAGTC	436	
QY	362	ACCAGAGGTGACCGTTTCGATGAGGGTGTCTCATTCGAGTTCACCGGTGTGCTT	421	
Db	437	ACCAGAGGTGACCGTTTCGATGAGGGTGTCTCATTCGAGTTCACCGGTGTGCTT	496	
QY	422	TCGTGCTCTACAAACGACCACTGATGTTGTTGCTTCTTACTGACCAACA	481	
Db	497	TCGTGATGTACAAACGACCACTGATGTTGTTGCTTCTTACTGACCAACA	556	
QY	482	ACAACGACCAACGAGTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAAC	541	
Db	557	ACAACGACCAACGAGTTGATCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAAC	616	
QY	542	AAGAGTCTTTAAGGTACCGAACCAAGAGCAGCAAAAGCAG	581	
Db	617	AAGAGTCTTTAAGTACCAACGACCAACGAGCAGCAAGAG	656	
RESULT 14				
BD175913				
LOCUS	BD175913 1446 bp DNA linear PAT 18-MAR-2003			
DEFINITION	Regulation of soybean seed protein gene of definite class.			
ACCESSION	BD175913			
VERSION	BD175913.1 GI:29121615			
KEYWORDS	JP 2002262894-A/13.			
SOURCE	unidentified			
ORGANISM	unidentified			

unclassified.
1 (bases 1 to 1446)
Kinney,A.J. and Fader,G.M.
Regulation of soybean seed protein gene of definite class
Patent: JP 2002262894-A 13 17-SEP-2002;
EI DU PONT DE NEMOURS AND CO
OS Unidentified
PN JP 2002262894-A/13
PD 17-SEP-2002
PF 28-FEB-2002 JP 2002053791
PI 14-JUN-1996 US 60/019940
PI ANTHONY J KINNEY,GARY MICHAEL FADER
PC C12N15/09,A01H5/00,C12N5/10,C12N5/00,C12N5/00 CC
Strandedness: Single;
CC Topology: linear;
CC Regulation of soybean seed protein gene of definite class FH
Key Location/Qualifiers
FT source 1..1446
FT location/Qualifiers
FT /organism='Unidentified'.
1..1446
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'
ORIGIN
Query Match 34.0%; Score 517.4; DB 6; Length 1446;
Best Local Similarity 62.8%; Pred. No. 3.3e-124;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;
QY 5 AGCAACCGGAGCAACGCGTCCAGTTCAGCGCCCTCAATGCCGACAGACCTGACAATC 64
DB 71 AGCAGCCACAGCAAAAGAGTGCAGATCCAAACCCCTCAATGCCCTAAACCGGATAACC 130
QY 65 GCATTGAATCAGAGGGCGGTTACATTGAGAGCTTGGAAACCCCAACCAACAGGAGTTCCGAAT 124
DB 131 GTATAGAGTCAGAGGTGGCTTCAATGAGACATGNAACCCCTAACCAAGCCATTCCAGT 190
QY 125 GCGCGGGTGCCTCTCTCGCTTAGTCTCCGCGGAGAGCCCTTCGTAGGCTTTCT 184
DB 191 GTGCGGTGTGCGCTCTCTCGCTGACCCCTCAACCGCAAGCGCCCTTCGAGAGCTTTCT 250
QY 185 ACTCAATGTCCTCCAGAGATCTTCATCCAGCAAGAGGAGGATCTTTGGGTGATAT 244
DB 251 ACACCAAGCTTCCCGAGAGATCTACATCCAAAGGTAGTGTATTTTGGCATGATAT 310
QY 245 TCCCTGCTTGTCTTAGACACTATGAAGAGCCTCAACACAGGTCGTCTCAGTCCC 304
DB 311 TCCCGGTGTCTAGCACATTGAAGAGCT----- 342
QY 305 AAGACCACCAAGACGTCTCCAGGAGAGAACCAAGCCCAACAGCAACGAGATAGTCAAC 364
DB 343 -----CAACAAAGAGCAAAAGCAGCAGCGCCCAAGAGCCGTCAAC 382
QY 365 AAGAGTGACCGTTTCGATGAGGCTGATCTCATGTCAGTTCCTCCACCGGTGTCCTTCT 424
DB 383 AGAAGATCTATCACTTCAGAGAGGATGATTTGATGTCAGTGCACACCGGTTTGCATACT 442
QY 425 GGCTCTCAACAGCAGCAGCACTGATGTTGCTGTTTCTCTTACTGACACCAACAAACA 484
DB 443 GGATGTACACATGAGACACTCTCTGTTGTCGCTTTCTTTATTGACACCAACAGCT 502
QY 485 ACACAACAGCTTGTATGATTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAAG 544
DB 503 TCCAGAACAGCTCGACAGATGCTAGGAGATTCTATCTTGTCTGGGAAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGTACAGCAACAGCAGACAAAGCAGACGAGAGCTTACCATATAGCC 604
DB 563 AGTTTCTACAGTATCACGCCACAGAGCAGCAAG----- 595
QY 605 CATACAGCCCGCAAGTGCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACAGC 664
DB 596 -----GAGGTACTC 604

QY 665 ACAGCCGACAGAAACGAGCAGGACACAGAGAAAGAAACGAGGTGGAAACATCTTCAGCG 724
DB 605 AAAGCCAGAAAGGAAAGCGCTCAGCAAGAAAGAAAGAAAGGAGCGCAGCATATTAGTG 664
QY 725 GCTTTCAGCGCGGAGTTCTCTGGAACAAGCCTTCCAGGTTGACGACAGACATAGTCAAA 784
DB 665 GCTTGGCCCGGAATTTCTTGAACATGCGTTC---GTCTGGACAGGCAATAGTGAGAA 721
QY 785 ACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGGCATTGTGACAGTGGAGGAGGCC 844
DB 722 AGCTCAAGGTGAGAAAGAGAGAGAGAGAGGAGGTCCTTGTGACAGTGAAGAGAGGTC 781
QY 845 TCAGATCTTGAGCCCGCCAGATAGAAAGAGAGCTGCCGACGAAAGAGGAGGAATACGATGAAG 904
DB 782 TCAGCGTGAATAAGCCCAACCCACCGNAGAGCAGCAACAAAGACCCGAGGAGAGGAGAGC 841
QY 905 ATGAATATGAATACGATGAAGAGGATAGAAGGCGTGCAGGGAAGCAGAGGCGAGGGGA 964
DB 842 CAGATTGTGACGAGAAAGACAAACAT-----TGCCTAAGGCCAAAGCAGAA 886
QY 965 ATGCTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTTGGTAGAAACAGAT 1024
DB 887 ATGCGATTGACGAGACCAATTTGCACAATGAGACTTCGCCACAACATTTGGCCAGACTTTCAT 946
QY 1025 CCCTGACATCTACAACCTCAAGCTGTGTCTACTCAAAAAGTGCACAGGATCTCAACCTTC 1084
DB 947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCCGCTACCCAGCTCTCGACTTC 1006
QY 1085 TAATACCTTAGGTGCTTGGACCTAGTGTGTAATATGMAATCTCTACAGGAATGCATTGT 1144
DB 1007 CAGCCCTCTGCTGCTCAAACTCAGTGCCTGCTTGTGATCTCCTCGAAGATGCTATGT 1066
QY 1145 TTGCTGCTCACTACACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC 1204
DB 1067 TCGTGCCACACTACAACCTGAACGCAAAACAGCATAATATACGCAATTGAATGGAGCGGCAT 1126
QY 1205 AGTGAAGTCTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC 1264
DB 1127 TGTACAAGTGGTGAATTCGCAATGCTGAGAGAGTGTGATGGAGAGCTGCAAGAGGGAC 1186
QY 1265 ACCTGCTTGTGGTGCACAGAACTTCGCGTCTCGTGAAGTCCAGAGCGGAACTTCG 1324
DB 1187 AGTGTAAATGTGCCCAAAACTTTGCGTGTGCTGCAAGTCAAGAGGACACACTTCG 1246
QY 1325 AATACGTGGCATTCAGACAGACTCAAGGCCACGACATAGCCAACTCGCCGCTGAACAACT 1384
DB 1247 AGTATGTTTCAATTCAGAGCAATGATAGACCTCGATCGCAACCTTGCAGGTGCAAACT 1306
QY 1385 CCGTCATAGATAACCTCGCGGAGGAGTGGTTGCAAAATTCATATGGCTCCAAAGGGAGC 1444
DB 1307 CATTGTTGAACGCAATTCGCGGAGGAAAGTGAATTCAGCAAACTTTAACTTAAGGAGGCGAGC 1366
QY 1445 AGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCTGTTTCCACCGCTCTCAGCAGT 1504
DB 1367 AGCCAGGAGGTCAGAAACAAACCCCTTTCAGTCTCTGTTTCCACTTAAGAGGAGTCTC 1426
QY 1505 CTCGAGGCGCTGTGGCTTA 1523
DB 1427 AGAGGAGAGTTGTGGCTTA 1445

RESULT 15
AR202578
LOCUS AR202578 1446 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 13 from patent US 6362399.
ACCESSION AR202578
VERSION AR202578.1 GI:20257117
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1446)

AUTHORS Kinney,A.John. and Fader,G.Michael.
TITLE Suppression of specific classes of soybean seed protein genes
JOURNAL Patent: US 6362399-A 13 26-MAR-2002;
FEATURES source Location/Qualifiers
1.1446
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 34.08; Score 517.4; DB 6; Length 1446;
Best Local Similarity 62.88; Pred. No. 3.3e-124;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;
QY 5 AGCAACCGGAGGAAACGCGTCCAGTTCAGGCGCTCAATGCGCAGAGACCTTGACAATC 64
DB 71 AGCAGCCACAGCAACAGAGTCCAGATCCAAAGCCCTAATGCGCAGGATAC 130
QY 65 GCATTGAATCAGAGGGCGGTTCATTGAGACTTTGGAAACCCCAACACAGAGAGTTGCAAT 124
DB 131 GTATAGAGTCAGAAAGTGGCTTCATTGAGACATGGAACCTTAACAACAGCCATTCCAGT 190
QY 125 GCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCGCTTCT 184
DB 191 GTGCGGCTGTGCGCTCTCTCGCTGACCTCAACCGCAACGCCCTTCGCGAGACCTTCCT 250
QY 185 ACTCCAAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTGATAT 244
DB 251 ACACCAACGCTCCCCAGGAGATCTACATCCCAACAGGTAGTGATATTTTGGCATGATAT 310
QY 245 TCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCACTGCC 304
DB 311 TCCGCGGTTCTCTAGCACATTTGAAGGCT----- 342
QY 305 AAAGACCACAGACGCTCTCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGTCACC 364
DB 343 -----CAACAAAGAGGACAAAGCAGCAGGCCCAAGACCGTCACC 382
QY 365 AGAAGGTGACCGTTCGATGAGGTGATCTCATTTGCAGTTCACCGGTGTTGCTTCT 424
DB 383 AGAAGATCTACTCTCAGAGAGGGTGTATTTGATGCGATGCCAACCGGTTTGCATACT 442
QY 425 GGCTCTCAACAGCACAGCACTGATGTTGCTGTTCTCTTACTGACACCAACAA 484
DB 443 GGATGTACAAACATGAAGACACTCTCTGTTGCGGTTCTCTATTGACACCAACAGCT 502
QY 485 ACACACACGAGTGTATCAGTTCACGAGGATTCATTTGGCTGGGACACAGGACGAG 544
DB 503 TCCAGAACCCAGCTCGACCATGCTAGGAGATTTCTATCTTGCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAGGAGCTTACCATATAGCC 604
DB 563 AGTTCTACAGTATCAGCCACAGACGACGAG----- 595
QY 605 CATACAGCCCGCAAGGTGACGCTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGACAGC 664
DB 596 -----GAGGTACTC 604
QY 665 ACAGCCGACAGACGAGCAGGACGAGCAAGAAACGAGGTGGAAACATCTTCAGCG 724
DB 605 AAAGCCAGAAAGGAAACGCGTCAGCAAGAAAGAAACGAAAGGAGGCGCATATTGAGTG 664
QY 725 GCTTCACGCGCGGAGTTCTTGAAACAAAGCCCTTCAGGTTTGACGACAGACAGATAGTCAAA 784
DB 665 GCTTCGCGCGCGAATCTTGGACATGCGTTC---GTCTGGGACAGGCGCATAGTGAGAA 721
QY 785 ACCTAAGAGCGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGAAGTGGAGGGAGGCC 844
DB 722 AGCTTACAAAGGTGAGAACGAAAGAGGAGAGAGGGGTGCCATTGTGACAGTGAAGAGGAGTC 781
QY 845 TCAGATCTTGGCCAGATAGAAAGAGAGCTGCCGACGAGAGAGGAGATACGATCAAG 904
DB 782 TCAGCGTGATAGGCCACCCACCGAGAGAGCAGCAACAAAGCCCGAGAGAGGAGAGAC 841

QY 905 ATGAATATGAATACGATGAAGAGGATAGAAAGGCGTGGCAGGGGAAAGCAGAGCGCAGGGGA 964
DB 842 CAGATTGTGACGAGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
QY 965 ATGGTATTGAAGAGAGAGTCTGCACCGCAAGTGTCTAAAAGAACATTTGTTAGNAACAGAT 1024
DB 887 ATGGCAATTGACGAGACCAATTTGCACAAATGAGACTTTGCCACAACTTGGCCAGACTTCAT 946
QY 1025 CCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAAATGCGCAACGATCTCAACCTTC 1084
DB 947 CACCTGACATCTCAACCTCAAGCTGGTAGATCAACCGCTACCAGCTTCGATTC 1006
QY 1085 TAATCTTAGTGGCTTGGACCTAGTGTGTAATA TGGAAATCTCTACAGGAATGCAATGT 1144
DB 1007 CAGCCCTCTCGTGGCTCAAACTCAGTGGCCAGTTTGGATCACTCCGCAAGAAATGCTATGT 1066
QY 1145 TTGTCGCTCACTACAACACCAACGACACAGCATCATATATATCGATTGAGGGGACGGGCTC 1204
DB 1067 TCGTGCACACTACAACTTGAACGCAACAGCATATATACGATTTGAATGGACGGGCAT 1126
QY 1205 AGTGCACAGTCTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC 1264
DB 1127 TGTACAGTGGTGAATTGCAATGGTAGAGAGTGTGATGGAGAGCTGCAAGAGGGAC 1186
QY 1265 AGTGTCTTGTGGTGCACAGAACTTTCGCGCTCGCTGGAAAGTCCCAGAGCGAGAACTTCG 1324
DB 1187 AGTGTGTTAATTGTGCCACAAAACCTTTGCGTGGCTGCAAGATCACAGAGCGCAACTTCG 1246
QY 1325 AATACGTGCAATTCAGACAGACTCAAGGCCAGACATAGCCAACTCGCGCGGTGAAAACCT 1384
DB 1247 AGTATGTTTTCATTCAAGACCAATGATAGACCTCGATCGGCAACCTTGCAGGTGCAAACT 1306
QY 1385 CCGTCATAGATAACCTGCGGAGGAGTGGTTGCAAAATTCATATGCGCTCCAAAGGGAGC 1444
DB 1307 CATTTGTTGAACGATTCGCGGAGGAGTGAATTCAGCAAACTTTTAACCTAAGGAGCAGC 1366
QY 1445 AGGCAAGGCGAGCTTAAGAAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGT 1504
DB 1367 AGGCCAGGCGAGTCAAGAACAAACACCCCTTTCAGCTTCCTGTTCCACCTAAGGAGTCTC 1426
QY 1505 CTCGAGGCGCTGTGCTTA 1523
DB 1427 AGAGGAGAGTTGTGGCTTA 1445

Search completed: August 24, 2005, 00:57:21
Job time : 4522.08 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 14:12:30 ; Search time 570.653 Seconds
(without alignments)
15809.409 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cggcagcaaccggaggagaa.....ctccgaggcgtgtggtcttaa 1524

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1524	100.0	1524	2	Az06384 Peanut al
2	1524	100.0	1524	2	Az22280 Ara h 3 a
3	1524	100.0	1524	8	Abx70612 Peanut Ar
4	1524	100.0	1524	10	Adg27543 Peanut CD
5	1520.8	99.8	1524	4	Aas08540 DNA encod
6	1391.2	91.3	1855	4	Aaf90338 Peanut al
7	1317.4	34.0	1446	2	Aav17567 Coding se
8	517.4	34.0	1446	6	Abss55193 Glycine m
9	517.4	34.0	1446	10	Adh89252 G. max gl
10	517.4	34.0	1446	12	Adg43987 G. max gl
11	513	33.7	1488	2	Aav17565 Coding se
12	513	33.7	1488	6	Abss55191 Glycine m
13	513	33.7	1488	10	Adh89244 G. max gl
14	513	33.7	1488	12	Adg43979 G. max gly
15	513	33.7	1743	3	Aaz92638 cDNA enco
16	505.6	33.2	1458	6	Abss55192 Glycine m
17	505	33.1	1746	4	Aad17536 Soybean g
18	504	33.1	1458	2	Aav17566 Coding se
19	504	33.1	1458	10	Adh89246 G. max gl
20	504	33.1	1458	12	Adg43981 G. max gl

21	502	32.9	1712	2	AAQ05359	Glycinin
22	498.6	32.7	1746	2	AAQ05360	Glycinin
23	224	14.7	1554	6	ABSS5194	Glycine m
24	198	13.0	1551	2	AAV17569	Coding se
25	198	13.0	1551	10	ADH89250	G. max gl
26	198	13.0	1551	12	ADG43985	G. max gl
27	198	13.0	1786	2	AAQ05357	Glycinin
28	193.2	12.7	3527	11	ADL90167	DNA encod
29	189	12.4	1374	12	ADQ43060	
30	188.8	12.4	1786	1	AAAN60939	Sequence
31	172.2	11.3	1689	10	ADH89248	G. max gl
32	172.2	11.3	1689	12	ADG43983	G. max gl
33	170.6	11.2	1899	2	AAQ05358	Glycinin
34	165.2	10.8	1646	1	AAH82246	Rice stor
35	162	10.6	1500	10	ADC08274	Rice DNA
36	153.2	10.1	1689	2	AAV17568	Coding se
37	153.2	10.1	1689	6	ABSS5195	Glycine m
38	151.8	10.0	1482	10	ADH89254	Sunflower
39	151.8	10.0	1482	12	ADG43989	H. annuus
40	149.8	9.8	1896	1	AAH60940	Sequence
41	147.8	9.7	1512	10	ADC07963	Rice DNA
42	144.2	9.5	2058	12	ADJ44956	Plant cDN
43	143	9.4	1512	10	ADC07957	Rice DNA
44	142	9.3	1706	2	AAH37335	Coffee et
45	140.2	9.2	1685	2	AAQ54819	Oat globi

ALIGNMENTS

RESULT 1	
AAZ06384	
ID	AAZ06384 standard; DNA; 1524 BP.
XX	
AC	AAZ06384;
XX	
DT	17-OCT-2003 (revised)
DT	09-NOV-1999 (first entry)
XX	
DE	Peanut allergen, Ara h 3.
XX	
KW	allergy; immune response; transgenic; allergen; epitope;
KW	immunoglobulin E; Ig E; binding site; peanut; ds.
XX	
OS	Arachis hypogaea.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1524
FT	/tag= a
FT	/product= "Ara h 3"
FT	/note= "First three amino acids not coded for"
XX	
PN	WO9938978-A1.
XX	
PD	05-AUG-1999.
XX	
PP	29-JAN-1999; 99WO-US002031.
XX	
PR	31-JAN-1998; 98US-0073283P.
PR	13-FEB-1998; 98US-0074590P.
PR	13-FEB-1998; 98US-0074624P.
PR	13-FEB-1998; 98US-0074633P.
PR	27-AUG-1998; 98US-00141220.
XX	
PA	(UYAR-) UNIV ARKANSAS.
PA	(UYNV) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.
PA	(SOSI/) SOSIN H.
XX	
PI	Sosin H, Bannon GA, Burks AW, Sampson HA;
DR	WPI; 1999-479189/40.
DR	P-PSDB; AAV15246.
XX	

PT Modified allergen with reduced IgE binding, useful for treating e.g. allergies.
XX
PS Disclosure; Page 38-39; 46pp; English.
XX
CC This is the nucleotide sequence of the Ara h 3 protein from Arachis hypogaea. The Ara h 3 protein has 4 IgE (Immunoglobulin E) binding epitopes, one of which is immunodominant (AAV15281). By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1524 BP; 455 A; 390 C; 398 G; 281 T; 0 U; 0 Other;
Query Match 100.0%; Score 1524; DB 2; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCAGCAACCGGAGGAGACGGGTCCAGTTCAGCGCCTCAATGCGCAGAGACCTGAC 60
DB 1 CGGCAGCAACCGGAGGAGACGGGTCCAGTTCAGCGCCTCAATGCGCAGAGACCTGAC 60
QY 61 AATCGCATTAATCAGAGGGCGGTACATTTGAGACTTGGAAACCCCAACACAGGAGTTC 120
DB 61 AATCGCATTAATCAGAGGGCGGTACATTTGAGACTTGGAAACCCCAACACAGGAGTTC 120
QY 121 GAATGCGCGCGGTGCGCCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180
DB 121 GAATGCGCGCGGTGCGCCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180
QY 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGATATTTGGGTTG 240
DB 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGATATTTGGGTTG 240
QY 241 ATATTCCTGCTGTTCTAGACACTATGAAGAGCCTCAACAGAGGTGCTGATCTCAG 300
DB 241 ATATTCCTGCTGTTCTAGACACTATGAAGAGCCTCAACAGAGGTGCTGATCTCAG 300
QY 301 TCCCAAGACACCAAGAGCTCTCCAGGAGAGAGCAAGCCCAACAGCAACAGAGATGT 360
DB 301 TCCCAAGACACCAAGAGCTCTCCAGGAGAGAGCAAGCCCAACAGCAACAGAGATGT 360
QY 361 CACCAGAGGTGACCGTTCGATGAGGGTGATCTCATTTGCACTTCCACCGGTGTTGCT 420
DB 361 CACCAGAGGTGACCGTTCGATGAGGGTGATCTCATTTGCACTTCCACCGGTGTTGCT 420
QY 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTTCTTACTGACCAAC 480
DB 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTTCTTACTGACCAAC 480
QY 481 AACACGACACCAAGCTTGTATCAGTTCCTCCAGGAGATTCATTTGGTGGAAACAGGAG 540
DB 481 AACACGACACCAAGCTTGTATCAGTTCCTCCAGGAGATTCATTTGGTGGAAACAGGAG 540
QY 541 CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
DB 541 CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
QY 601 AGCCCATACAGCCGCGAAAGTCAAGCTAGACAGAAAGAGCGGTGAATTTAGCCCTCAGGA 660
DB 601 AGCCCATACAGCCGCGAAAGTCAAGCTAGACAGAAAGAGCGGTGAATTTAGCCCTCAGGA 660
QY 661 CAGCAGACCGCGCAGAGAAACGAGCAGACAGAGAAAGAAACGAAGTGGAAACATCTTC 720
DB 661 CAGCAGACCGCGCAGAGAAACGAGCAGACAGAGAAAGAAACGAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTACCGCGGAGTTCCTGGAAACAAAGCTTCAGGTTGACGACAGACAGATAGTG 780
DB 721 AGCGGCTTACCGCGGAGTTCCTGGAAACAAAGCTTCAGGTTGACGACAGACAGATAGTG 780
QY 781 CAAAACTTAAGCGCGAGACCGGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
DB 781 CAAAACTTAAGCGCGAGACCGGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840

DB 781 CAAAACTTAAGCGCGAGACCGGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
QY 841 GGCCTCAGAAATCTTGAGCCAGATAGAAAGAGAGAGCGTCCGAGCAAGAGAGGAATACGAT 900
DB 841 GGCCTCAGAAATCTTGAGCCAGATAGAAAGAGAGAGCGTCCGAGCAAGAGAGGAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGCGGAGAGAGGAGCAGG 960
DB 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGCGGAGAGAGGAGCAGG 960
QY 961 GGAATGATGATGAAGAGAGAGTCTGTCACCGCAAGTCTTAAAGAAACATTTGGTAGAAGC 1020
DB 961 GGAATGATGATGAAGAGAGAGTCTGTCACCGCAAGTCTTAAAGAAACATTTGGTAGAAGC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAAC 1080
DB 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAAC 1080
QY 1081 CTTCTAATACTTAGGTGGCTTGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
DB 1081 CTTCTAATACTTAGGTGGCTTGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTTGTGCTCATAACACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
DB 1141 TTGTTTGTGCTCATAACACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
QY 1201 GCTCAGTCCAAAGTCTGTCAGACAGCAACGCAACAGAGTGTACGAGAGAGCTTCAAGAG 1260
DB 1201 GCTCAGTCCAAAGTCTGTCAGACAGCAACGCAACAGAGTGTACGAGAGAGCTTCAAGAG 1260
QY 1261 GGTCAAGTCTTGTGTGTGTCACAGAACTTCGCGCGTTCGCTGGAAGTCCAGAGCGAGAAC 1320
DB 1261 GGTCAAGTCTTGTGTGTGTCACAGAACTTCGCGCGTTCGCTGGAAGTCCAGAGCGAGAAC 1320
QY 1321 TTGCAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCACCTCCCGGTGAA 1380
DB 1321 TTGCAATAGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCACCTCCCGGTGAA 1380
QY 1381 AACTCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGG 1440
DB 1381 AACTCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGG 1440
QY 1441 GAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTCTTCCGTCACCGTCTCAG 1500
DB 1441 GAGCAGGCAAGCAGCTTAAGAAACAAACCCCTTCAAGTCTTCCGTCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
DB 1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
RESULT 2
AAZ22280 standard; DNA; 1524 BP.
ID AAZ22280
XX AAZ22280;
AC AC
XX 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX
DE Ara h 3 allergen encoding DNA.
KW Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
XX allergic reaction; ss.
OS Arachis hypogaea.
XX
FH Key Location/Qualifiers
FT 1. .1524
FT /*tag= a
FT /transl_except= (pos:490..491, aa:Glu, Phe)
FT /note= "there is an apparent 6 nucleotide deletion which
FT encodes for residues Glu and Phe respectively"

CC previously displayed allergic symptoms when exposed to the antigen.
 CC Following this an antigen-specific IgE present on one or more mast cells
 CC or basophils in the individual's serum is identified. The individual is
 CC then contacted with a peptide corresponding to a portion of the antigen,
 CC which is selected, formulated, and delivered so that binding of the
 CC peptide to antigen-specific IgE is reduced as compared with IgE binding
 CC of intact antigen. The composition is also useful for treating and
 CC preventing allergic reactions
 XX

Query Match 99.8%; Score 1520.8; DB 4; Length 1524;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1522; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGG CAG CAA CCG GAG GAG AAG CGG TGC CAG TTC CAG CGC CTCAAT GCG CAG ACG CTG AC 60
 Db 1 CGG CAG CAA CCG GAG GAG AAG CGG TGC CAG TTC CAG CGC CTCAAT GCG CAG ACG CTG AC 60

Qy 61 AAT CGC ATT GAA TCA GAG GGG GGT TAC ATT GAG ACT TGG AAC CCG CCA CCA CCG AGG ATT C 120
 Db 61 AAT CGC ATT GAA TCA GAG GGG GGT TAC ATT GAG ACT TGG AAC CCG CCA CCA CCG AGG ATT C 120

Qy 121 GAAT GCG CGG CGT GCG CCT CTCT CGCT TAG TCT CCG CGC GAC CAG CGC CTTCG TAG GCCT 180
 Db 121 GAAT GCG CGG CGT GCG CCT CTCT CGCT TAG TCT CCG CGC GAC CAG CGC CTTCG TAG GCCT 180

Qy 181 TTCT ACT CCA ATG TCT CCCC AGG AGAT CTTCAT CAG CAA GAA GGG GAT CTT TGG GTTG 240
 Db 181 TTCT ACT CCA ATG TCT CCCC AGG AGAT CTTCAT CAG CAA GAA GGG GAT CTT TGG GTTG 240

Qy 241 ATATT CCT TGG TGT CTTAG ACATAT GAA GAG CTT CAC A CAG AGT CGT CAG ATCT CAG 300
 Db 241 ATATT CCT TGG TGT CTTAG ACATAT GAA GAG CTT CAC A CAG AGT CGT CAG ATCT CAG 300

Qy 301 TCC CAA AGA CAC CAA GAG CTT CCA AGG AAG AAG CAA AGC CAA CAG CAG AAG TAGT 360
 Db 301 TCC CAA AGA CAC CAA GAG CTT CCA AGG AAG AAG CAA AGC CAA CAG CAG AAG TAGT 360

Qy 361 CAC CAG AAG GTG CAC CGT TCG ATG AGG GTG ATCTCA TTG CAG TTTCC ACC CGG TTTGCT 420
 Db 361 CAC CAG AAG GTG CAC CGT TCG ATG AGG GTG ATCTCA TTG CAG TTTCC ACC CGG TTTGCT 420

Qy 421 TTCT GGT CTTACA CAG CAG CAG CAG CAG TGTG TGTG TGTG TGTG TGTG TGTG TGTG TGTG 480
 Db 421 TTCT GGT CTTACA CAG CAG CAG CAG CAG TGTG TGTG TGTG TGTG TGTG TGTG TGTG TGTG 480

Qy 481 AAC AAG CAA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 540
 Db 481 AAC AAG CAA CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 540

Qy 541 CAAG AGT TTTAAG GTTAC CAG CAA CAA GAG CAG CAA GAG CAG CAA GAG CAG CAA GAG CAG 600
 Db 541 CAAG AGT TTTAAG GTTAC CAG CAA CAA GAG CAG CAA GAG CAG CAA GAG CAG CAA GAG CAG 600

Qy 601 AGCC CATA CAG CCG CAA AGT CAG CTTAG A CAG AAG AGC GTG AAT TTTAG CCG CTTAG GGA 660
 Db 601 AGCC CATA CAG CCG CAA AGT CAG CTTAG A CAG AAG AGC GTG AAT TTTAG CCG CTTAG GGA 660

Qy 661 CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 720
 Db 661 CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 720

Qy 721 AGCG GCTT CAG CCG GAG TTTCTT GGA CAA GAG CTTCC AGG TTTG A CAG CAG CAG A GATG 780
 Db 721 AGCG GCTT CAG CCG GAG TTTCTT GGA CAA GAG CTTCC AGG TTTG A CAG CAG CAG A GATG 780

Qy 781 CAAA A CTTAAG CCG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 840
 Db 781 CAAA A CTTAAG CCG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG CAG 840

Qy 841 GGCCT CAG AATCTT GAG CCG CAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 900
 Db 841 GGCCT CAG AATCTT GAG CCG CAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 900

Db 841 GGCCT CAG AATCTT GAG CCG CAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG AAG 900

Qy 901 GAAG ATGA ATATGA ATACG ATGAAG AGG ATGAAG AGG CTGG CAG GGAAG CAG AGG CAGG 960

Db 901 GAAG ATGA ATATGA ATACG ATGAAG AGG ATGAAG AGG CTGG CAG GGAAG CAG AGG CAGG 960

Qy 961 GGAATGGTATTGAAGAGAGAGATCTCCACCGCAAGTGTCTAAAGAGACATTTGGTAGAATC 1020

Db 961 GGAATGGTATTGAAGAGAGAGATCTCCACCGCAAGTGTCTAAAGAGACATTTGGTAGAATC 1020

Qy 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCAACGATCTCAAC 1080

Db 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCAACGATCTCAAC 1080

Qy 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGCTGAATATGGAATCTCTACAGGAATGCA 1140

Db 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTCTGCTGAATATGGAATCTCTACAGGAATGCA 1140

Qy 1141 TTGTTTGTGCTCAGTACACCAACGACACAGCATCATATATCGATTGAGGGGACGG 1200

Db 1141 TTGTTTGTGCTCAGTACACCAACGACACAGCATCATATATCGATTGAGGGGACGG 1200

Qy 1201 GCTCAGTGCAGTCCGTGGACAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260

Db 1201 GCTCAGTGCAGTCCGTGGACAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260

Qy 1261 GGTCAAGTCTTGTGTGCGCACAGAACTTCGCGCTCGCTGGAAAGTCCCAAGAGCGAAG 1320

Db 1261 GGTCAAGTCTTGTGTGCGCACAGAACTTCGCGCTCGCTGGAAAGTCCCAAGAGCGAAG 1320

Qy 1321 TTGGAATAGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380

Db 1321 TTGGAATAGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380

Qy 1381 AACTCCGTCAATAGATAACCTCGCGAGGAGTGTGCAAAATTCATATGCGCTCCAAAGG 1440

Db 1381 AACTCCGTCAATAGATAACCTCGCGAGGAGTGTGCAAAATTCATATGCGCTCCAAAGG 1440

Qy 1441 GAGCAGGCAAGGAGCTTAAGAACAAACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500

Db 1441 GAGCAGGCAAGGAGCTTAAGAACAAACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500

Qy 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524

Db 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524

RESULT 6
 ID AAF90338 standard; cDNA; 1855 BP.
 XX
 AC AAF90338;
 XX
 DT 11-SEP-2003 (revised)
 DT 23-JUL-2001 (first entry)
 XX
 DE Peanut allergen Ara h4 cDNA.
 XX
 KW Peanut; allergen; Ara h2; Ara h3; transgenic plant; allergy; ss.
 XX
 OS Arachis hypogaea.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..1593
 FT /*tag= a
 FT misc_feature 430..1190
 FT /*tag= b
 FT /note= "PCR amplified fragment"
 XX
 PN W0200136621-A2.
 XX
 PD 25-MAY-2001.
 XX

QY	1385	CCGTCATAGATAACCTCGCGAGAGAGTGGTTGCAAAATTCATATGGCTCCAAAGGGAGC	1444
Db	1307	CAATTGTTGAACGCATTGCGGAGGAAGTGATTTCAGCAAACTTTTAACTTAAGGAGGCAGC	1366
QY	1445	AGGCAAGCGCAGCTTAAAGAACAAACCCCTTTCAGTTCCTCGTTCCACCGTCTTCAGCAGT	1504
Db	1367	AGGCCAGGCGAGTCAAGAACAAACCCCTTTCAGCTTCCTCGTTCCACCTTAAGGAGTCTC	1426
QY	1505	CTCCGAGGGCTGTGGCTTA	1523
Db	1427	AGAGGAGAGTTGTGGCTTA	1445
RESULT 8			
ID	ABS55193	standard; DNA; 1446 BP.	
XX	XX	ABS55193;	
XX	DT	17-DEC-2002 (first entry)	
XX	DE	Glycine max (Soybean) var. Dare gene.	
XX	KW	Soybean; Glycinin; atomic coordinate data; processability; soya protein;	
XX	KW	Dare; protein co-ordinate data; gene; ds.	
XX	OS	Glycine max.	
XX	XX		
XX	Key	Location/Qualifiers	
XX	CD	1. 1446	
XX	FT	/*tag= a	
XX	FT	/product= "Glycine max (Soybean) var. Dare protein"	
XX	PN	JP2002193996-A.	
XX	XX		
XX	PD	10-JUL-2002.	
XX	XX		
XX	PF	21-DEC-2000; 2000JP-00405097.	
XX	XX		
XX	PR	21-DEC-2000; 2000JP-00405097.	
XX	XX	(KYOU) UNIV KYOTO.	
XX	XX		
XX	DR	WPI; 2002-685438/74.	
XX	DR	P-PSDB; ABG71266.	
XX	XX		
XX	PT	Glycinin, beta-conglycinin and proglycinin, their crystal structures, three dimensional coordinates, three dimensional structured and models and their uses.	
XX	XX		
XX	PS	Disclosure; Page 1271-1273; 1298pp; Japanese.	
XX	XX		
XX	CC	The present invention relates to a new Glycinin characterised by the atomic coordinate data fully defined in the specification. The structure can be used for improving processability of soya protein. The present nucleic acid sequence encodes the Glycine max (Soybean) var. Dare protein, as described in the specification	
XX	XX		
SQ	Sequence 1446 BP; 426 A; 361 C; 354 G; 305 T; 0 U; 0 Other;		
Query Match 34.0%; Score 517.4; DB 6; Length 1446;			
Best Local Similarity 62.8%; Pred. No. 1.4e-137;			
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;			
QY	5	AGCAACGGAGGAGAACGGGTGCCAGTTCAGCGCTCAATGCGCAGAGCTGACATC	64
Db	71	AGCAGCCACAGCAAAACGAGTGCAGATCCAAAGCCTCAATGCGCTTAAACCGGATAAC	130
QY	65	GCATTGATCAGAGGGGGTTTACATTGAGACTTGGAAACCCCAACACGAGGTTGGAAT	124
Db	131	GTATAGAGTCAGAGGTGGCTTCATTGAGACATGGAAACCCCTTAACAAAGCCATTCCAGT	190

QY	125	GCGCGGGTGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGCCCTTCT	184
Db	191	GTGCGGTGTTGCCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTCGACAGCTTCT	250
QY	185	ACTCAATGCTCCCAAGGAGATCTTATCCACAGCAAGAGGGGATATCTTTGGTTCATAT	244
Db	251	ACACCAACGCTCCCGAGGAGATCTATCCAAACAGGTAGTGGTATTTTGGCATGATAT	310
QY	245	TCCTCTGTTGCTTAGACACTATGAAGCCCTCACACACAAGGTGCTCGATCTCAGTCCC	304
Db	311	TCCGGGTGCTCTAGCACATTTGAAGCCT-----	342
QY	305	AAAGACCACCAAGAGCTCTCCAAGGAGAGACCAAGCCCAACAGCAACAGATAGTCAAC	364
Db	343	-----ChAACAAAGGACAAAGCAGCAGGCCCCCAAGACCGTCAAC	382
QY	365	AGAAGTGCAACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCACACCGGTGTTGCTTCT	424
Db	383	AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTGCGATGCGCAACCGGTTTGCATACT	442
QY	425	GGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAACA	484
Db	443	GGATGTACAAACAAAGAACACTCTCTGTTGTTGCGGTTTCTCTTATGTGACCAACAGCT	502
QY	485	ACGACAACCCAGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAG	544
Db	503	TCCAGAACCCAGCTCGACAGAGTCTATCTCTGCTGGGAACCAAGAGCAAG	562
QY	545	AGTTCTTAAGGTACAGCAACAAACGACAAAGCAGACAGCAAGAAAGCTTACCATATAGCC	604
Db	563	AGTTTCTCAGTATCAGCCACAGAGCAGCAAG-----	595
QY	605	CATACAGCCGCAAGTTCAGCTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGACAGC	664
Db	596	-----GAGGTACTC	604
QY	665	ACAGCCGACAGAACGACGAGCAAGAAAGAAAGAAACGAAAGTGGAAACATCTTTCAGCG	724
Db	605	AAAGCCAGAAAGGAAAGCGTCAGCAAGAAAGAAAGAAACGAAAGGAGCAGCATATTGAGTG	664
QY	725	GCTTCACGCGGAGTTCCTGGAAACAGCTTCCAGGTTGACACACAGACAGATAGTCAAA	784
Db	665	GCTTCGCGCGGAATTTCTTGGAAACATGCGTTC---GTGCTGGACAGGCAAGTAGTGAGAA	721
QY	785	ACCTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGGAGCCATTTGTGACGTGAGGGGAGGCC	844
Db	722	AGCTACAGGTGAGAACGAGAGGAGAGAGAGGGTGCCATTTGTGACGTGAAGAGGGTCT	781
QY	845	TCAGAAATCTTGAGCCCAAGATAGAAAGAGACGTCGCCGACGAAGAGAGGAATACGATGAAG	904
Db	782	TCAGCGTGATAAGCCACCCACGGAAGAGCAGCAACAAAGACCCCGAGGAAGAGGAAGC	841
QY	905	ATGAATATGAATACGATGAGAGGATAGAGCGGTGGCAGGGGAAAGCAGAGCAGGGGGA	964
Db	842	CAGATTGTGACAGAAAGACAAACAT-----TSCCAAAAGCCAAAGCAGAA	886
QY	965	ATGGTATTGAAGAGACGATCTGACGCGAAGTGCTTAAAGAAACATTTGTTAGAAACAGAT	1024
Db	887	ATGGCATTGACGAGACCATTTGCACATGAGACTTCGCCACAACTTGGCCAGACTTCTAT	946
QY	1025	CCCCTGACATCTACAAACCTCAAGCTGGTTCTCAAAAATCGCCCAACGATCTCAACCTTC	1084
Db	947	CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCACAAACCGCTACCAACCTCGACTTC	1006
QY	1085	TAATACTTAGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGATGCAATGT	1144
Db	1007	CAGCCCTCTCGTGGCTCAAACTCAGTGGCCAGTTTGGATCACTCCGCAAGATGCTATGT	1066
QY	1145	TTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC	1204
Db	1067	TGTTGCCACATACAACTGAACCGCAACAGCATTAATATACGATTTGAATGACGGGCAT	1126
QY	1205	ACGTGCAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC	1264

Db 1127 TGGTACAGTGGTGAATTCGAATGGTGAGAGAGTGTGGTGGAGAGCTGCAGAGGGAC 1186
 Qy 1265 ACCTGCTTGTGGTCCACAGAACTTTCGCGCTCGCTGGAAAGTCCACAGAGGAGAACTTCG 1324
 Db 1187 AGGTGTTAAATGTGCCCAAACTTTGCGGTGGCTGCAAGATCACAGAGCGCAAACTTCG 1246
 Qy 1325 AATACGTGGCAATTCAGACAGACTCAAGGCCACATAGCCACCTCGCCGGTGCAAACT 1384
 Db 1247 AGTATGTTTCATTCAGACCAATGATAGACCTCGATCGGCAACTTTCAGAGTGCAAACT 1306
 Qy 1385 CCCTCATAGATAAATCTCCGAGGAGGTGGTGTGCAAAATTCATATGGCTCCAAAGGGAGC 1444
 Db 1307 CATTTGTTGAAGCATTCGCGGAGAGTGAATTCAGCAAACTTTAACTAAGGAGGAGC 1366
 Qy 1445 AGCAAGGAGCTTAAGAAACAAACCCCTTCAAGTTCCTTGGTTCACCGTCTCAGCAGT 1504
 Db 1367 AGGCAGGAGGCTCAAGAAACAAACCCCTTTCAGCTTCCTGGTTCACCTAAGGAGTCTC 1426
 Qy 1505 CTCGAGGGCTGTGGCTTA 1523
 Db 1427 AGAGGAGTGTGGCTTA 1445

RESULT 9

ADH89252
 ID ADH89252 standard; DNA; 1446 BP.

AC ADH89252;

XX AC

XX 06-MAY-2004 (first entry)

DT XX

DE XX

DE XX G. max glycinin subunit G3 DNA.

XX ds; gene; double stranded RNA; storage protein; 2S-albumen; 7S-globulin;

KW 11S/12S-globulin; zein-prolamine; homogenitistate metabolic pathway;

KW pharmaceutical; plant; abiotic stress; fatty acid composition;

KW lipid composition; oil composition; carbohydrate composition; colour;

KW pigmentation; pathogen resistance; fruit ripening delay; aging;

KW male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;

KW caffeine; theophylline; threonine biosynthesis; glycine.

XX Glycine max.

XX WO2003078629-A1.

XX 25-SEP-2003.

XX 17-MAR-2003; 2003WO-EP002735.

XX 20-MAR-2002; 2002DE-01012892.

PR (BADI) BASF PLANT SCI GMBH.

XX Kock M, Bauer J;

XX WPI; 2003-803889/75.

DR P-PSDB; ADH89253.

XX Reducing expression of at least two target genes, useful e.g. for

XX producing transgenic plants, using partly double-stranded interfering

XX RNA.

XX Claim 7; SEQ ID NO 27; 228pp; German.

XX This invention describes a novel method for reducing the expression of at

XX least two different endogenous target genes in a eukaryotic cell or

XX organism by introducing an RNA molecule that is at least partly double

XX stranded. The transcribed RNAs from at least two target genes have

XX homology below 90% and the RNA molecule is formed as a single, self-

XX complementary molecule. At least one of the double-stranded structures

XX formed from individual sense sequences has an even number of repeats of

XX 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At

CC least two target genes are selected from different classes of storage
 CC protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-
 CC prolamine and at least one of the sense sequences is identical to storage
 CC protein sequences or genes in the homogenitistate metabolic pathway or
 CC enzyme types, e.g. acetyl transferases, thioesterases, (de)branching
 CC enzymes or cellulases. The RNA of the invention, also related cassettes,
 CC expression systems, vectors and transgenic organisms are used for
 CC preparation of pharmaceuticals, in biotechnological processes and plant
 CC biotechnology, specifically in plants to improve protection against
 CC abiotic stress, to modify composition and/or content of fatty acids,
 CC lipids and oils, to modify carbohydrate composition, to alter colour or
 CC pigmentation, to reduce content of storage proteins, to increase
 CC resistance to pathogens, to inhibit stem break, to delay fruit ripening
 CC or aging, to induce male sterility, to reduce content of toxic or
 CC unwanted components, to modify lignification and/or lignin content, to
 CC modify the fibre component in foods or fibre quality in cotton, to reduce
 CC susceptibility to shock, to increase synthesis of Vitamin E, to reduce
 CC contents of nicotine, caffeine or theophylline and to increase methionine
 CC content, by reducing threonine biosynthesis. The method provides a rapid
 CC and efficient way of reducing gene expression, can inhibit more than one
 CC target gene, prevents development of multiple phenotypes (since the
 CC transcription rate is the same for all RNA sequences, significantly
 CC reducing the selection process required to produce an organism with
 CC effective suppression of all target genes), avoids problems of epigenic
 CC gene silencing, does not require synthesis of individual RNA sequences
 CC and the method can be applied to plants with complex (polyploid) genomes.
 CC No interference between the individual RNA sequences occur. This sequence
 CC represents a target gene used in the method of the invention.
 XX
 SQ Sequence 1446 BP; 426 A; 361 C; 354 G; 305 T; 0 U; 0 Other;

Query Match 34.0%; Score 517.4; DB 10; Length 1446;

Best Local Similarity 62.8%; Pred. No. 1.4e-137;

Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

Qy 5 AGCAACCGGAGGAGAACGGGTGCGAGTCCAGCGCTCAATGGCGAGAGACCTGCACATC 64
 Db 71 AGCAGCCACAGCAAAACGAGTGCCAGATCCACGCTCAATGCCCTTAAACCGGATACC 130
 Qy 65 GCATTCAATCAGAGCGGTTCATTGAGACTTGGAAACCCCAACACAGGAGTTGGAAT 124
 Db 131 GTATAGAGTCAGAAGGTGGCTTCATTGAGACATGGAACCTTAACAACAGCCATTCAGT 190
 Qy 125 GCGCCGGCTGCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCGCTTCT 184
 Db 191 GTGCGGTGTGCGCTCTCTCGCTGCAACCTCAACCGCAACGCCCTTCGAGAGCTTCT 250
 Qy 185 ACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTGTGATAT 244
 Db 251 ACACCAACGCTCCCGAGGAGATCTATCCACAGGTAGTGTATTTTGGCATGATAT 310
 Qy 245 TCCCTGGTTGTCTTAGACACTATGAAGAGCCTCACACACAAGGTCGTGATCTCAGTCCC 304
 Db 311 TCCCGGTGTCTTAGACACTATGAAGAGCCT----- 342
 Qy 305 AAAGACCAACAGAGCGTCTCAAGGAGAGAGACCAAGCAACAGCAACAGAGATAGTACC 364
 Db 343 -----CAACAAAAGGACAAAGAGCAGCGCCCAAGACCGTCAAC 382
 Qy 365 AGAAGGTGCACCGTTTCGATGAGGGTGATCTCAATTCAGTTCCACCGGTGTTGCTTCT 424
 Db 383 AGAAGATCTATCACTTCAGAGAGGGGTGATTTGATTCAGTGCACCGGTGTTGCTACT 442
 Qy 425 GGCTCTACAACGACGACACTGTGTTGTTGCTGTTTCTTCTTACTGACCAACAACA 484
 Db 443 GGATGTACAACATGAAGACACTCTCTGTTGTTGCTGTTTCTTATGACCAACAGCT 502
 Qy 485 ACAGCAACGAGCTTGTATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGAGCAAG 544
 Db 503 TCCAGAACCGAGCTCGACAGATGCTCTAGGAGATTTCTATCTTGTGGGAACCAAGAGCAAG 562
 Qy 545 AGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAGAGAGCTTACCATATAGCC 604

Query Match		33.7%;	Score 513;	DB 2;	Length 1488;
Best Local Similarity		62.7%;	Pred. No. 2.6e-136;		
Matches 968;		Conservative	0;	Mismatches 425;	Indels 150; Gaps 5;
QY	5	AGCAACGGAGGAGACGCGTGCAGTTCAGCGCCCTCAATGCGCAGAGACCTGACAATC	64		
DB	71	AGCAGCCCTCAGCAAAACGAGTGCAGATCAAAAACCTCAATGCCCTCAAAACGGGATAACC	130		
QY	65	GCATTGATCAGAGGGCGGTTCATTCAGACTTGGAAACCCCAACACACAGGAGTTCGAAT	124		
DB	131	GTATAGAGTCAAGAGGAGGCGCTCATTCAGACATGGAACCCCTAACAACAGCCATTCAGT	190		
QY	125	GCGCGCGGTGCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGAGCCCTTCT	184		
DB	191	GTGCGGTGTGCGCCCTCTCTCGTGCACCTCAACCGCAACGCGCTTCGTAGACCTTCCT	250		
QY	185	ACTCCAAATGTCGCCAGGAGATCTTCATCCAGAAAGAGGATATCTTTGGTTCATAT	244		
DB	251	ACACCAACGGTCCCGCAGGAATCTACATCCAAACAGGTAAAGGTATTTTGGCATGATAT	310		
QY	245	TCCCTGGTTCCTAGACACTATGAAGAGCCTCACACAGGTCTCGATCTCAGTCCC	304		
DB	311	ACCGGGTTCCTAGACACTTTGAAGAGCCTCA-----	344		
QY	305	AAAGACCACCAAGAGGTCTCCAGGAGAGAACCAAGCCAAAGCAGACGAGATAGTCACC	364		
DB	345	-----ACAACTCAACAAGAGACAAAGCAGCAGACACACAGACCGTCACC	391		
QY	365	AGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCACCGGTGTTGCTTTCT	424		
DB	392	AGAAGATCTATACTTCAGAGAGGGTGATTTGATCGCAGTGCCTACTGGGTGTTGCATGGT	451		
QY	425	GGCTCTACAGCAGCAGCAGCAGTGTGTTGCTGTTCTTCTACTGACACCAACAACA	484		
DB	452	GSATGTACAAATGAAGACACTCCCTGTTGTTGCCGTTCTATTATTGACACCAACAGCT	511		
QY	485	ACGACCAACCACTGTGATTCAGTTCGCCAGGAGATTCATTTGGCTGGGAACACCGGCAAG	544		
DB	512	TGGAGAACCACTCGACCATGCTTAGGATTTCTATCTTGTGGGACCAAGAGCAG	571		
QY	545	AGTTCTTAAGGTCAGCAACAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC	604		
DB	572	AGTTTCTAAATATCAGCAAG-----	594		
QY	605	CATACAGCCCGCAAGCTAGACACGAGCGTGAATTTAGCCCTCGAGGACAGC	664		
DB	595	-----CAAGGAGGTGATC-----	607		
QY	665	ACAGCCGAGAGACGAGCAGGACGAAGAAAGAAACGAAGGTGGAAACATCTTCAGCG	724		
DB	608	AAAGCCAGAAGGAAAGCATCAGCAGAGAGAGAAACGAGGAGCGCATATTGAGTG	667		
QY	725	GCTTCACGCGGAGTTCCTGGAAACAAGCCTTCAGGTTCAGCAGACAGATAGTCAAA	784		
DB	668	GCTTCACCTGGAATCTTGGAACTGATTCAGCGT---GGCAAGCAGATAGCGAAAA	724		
QY	785	ACCTAAGAGCGAGACCGAGAGTGAAGAGAGGAGGACCATTTGACAGTGGAGGGAGGCC	844		
DB	725	ACCTAAGAGAGAGACGAAGGGGAGACAAAGGAGGACCATTTGACAGTGAAGAGGAGT	784		
QY	845	TCAGAACTTTGAGCCCA-----GATAGAAAGAGACGTGCGCGACGAAGAGAGGAAT	895		
DB	785	TGAGCGTGATAAACCACCCACGAGCAGCAGCAACAAAGACCCCGAGAGAGAGAGAG	844		
QY	896	ACGATGAAGATGAATA-----TGAATACGATGAAGAGGATAGAAGCGGTG	940		
DB	845	AAGAAGAGGATGAGAAGGCCACAGTGCAGAGGTTAAAGCAAACTGCAACGCCCCGAG	904		
QY	941	GCAGGGAGCAGAGCGGCGGGAATGTTATTGAAGAGCATCTGCACCGCAAGTGCTA	1000		
DB	905	GAAGCCAAAGCAAGCAAGAAATGGCATTCACGAGACCATATGCACCATGAGACTTC	964		

QY	1001	AAAAGAAACATTGGTAGAAAACAGATCCCTGACATCTTACAACCCCTCAAGCTGGTTCACTCA	1060
DB	965	GCACAAACATTGGCCAGACTTTCATCACCTGACATCTTACAACCCCTCAAGCGGTAGGTCA	1024
QY	1061	AAACTGCCAAACATCTCAACCTTCTAACTACTTAGTGGTGGACCTAGTCTGTAATATG	1120
DB	1025	CAACGCCCAACAGCCTTGACTTCCAGGCCCTCTCGTGGCTCAGACTCAGTCTGAGTTTG	1084
QY	1121	GAAATCTCTACAGGAATGCAATTTGTCGCTCACTTACAACCAACCAACGACACAGCATCA	1180
DB	1085	GATCTCTCGCAAGAAATGCAATTTGTCGTCACACTTACAACCTGAACGCGAAGCATTA	1144
QY	1181	TATATCGATTGAGGGACGGGCTCACGTGCAAGTCTGTGACAGCAACCGCAACAGAGTGT	1240
DB	1145	TATAGCAATTGAATGGACGGGCAATTGATACAAGTGTGTGAATTTGCAACGGTGAGAGAGTGT	1204
QY	1241	ACGACGAGAGCTTCAAGAGGGTCAAGTCTTGTGTTGGTGCACAGAACTTCGCGGTGCTG	1300
DB	1205	TTGATGGAGAGCTGCAAGAGGGGACGGGTCTGATCGTGCACAAAACCTTTGTGGTGGCTG	1264
QY	1301	GAAAGTCCAGAGCAGAGAACTTCGAATACGTGSCATTTCAAGACAGACTCAAGGCCCAGCA	1360
DB	1265	CAAGATCAAGAGTGAACAATTCGAGTATGTGTCTTCAAGACCAATGATACACCCATGA	1324
QY	1361	TAGCCAAACCTCGCCGCTGAAAACTCCGTCTATAGATAAACCTGCGGAGGAGGTGGTTCAA	1420
DB	1325	TCGGCACTCTTCAGGGGGCAAACTCATTTGTGACGCAATTCACAGAGGAAGTGAATTCAGC	1384
QY	1421	ATTCTATATGGCTTCCAAAGGGAGCAGGCAAGCAGCTTAAAGAAACAACACCCCTTCAAGT	1480
DB	1385	ACACTTTCAACCTAAAAAGCCAGCAGGCGCAGAGATAAAGAAACAACAACCCCTTCAAGT	1444
QY	1481	TCCTGTTCCACCGCTCTCAGCAGTCTCGAGGCGCTGTGGCTTA	1523
DB	1445	TCTGTTCCACCTCAGGAGTCTCAGAGAGAGCTGTGGCTTA	1487
RESULT 12			
ABSS5191	ID ABSS5191 standard; DNA; 1488 BP.		
XX	AC	ABSS5191;	
XX	DT	17-DEC-2002 (first entry)	
XX	DE	Glycine max (Soybean) var. Shiotsurunoko gene #1.	
XX	KW	Soybean; Glycinin; atomic coordinate data; processability; soya protein; Shiotsurunoko; protein co-ordinate data; gene; ds.	
XX	OS	Glycine max.	
XX	FT	Key	Location/Qualifiers
XX	FT	CDS	1..1488
XX	FT		/tag= a
XX	FT		/product= "Glycine max (Soybean) var. Shiotsurunoko
XX	FT		protein #1"
XX	PN	JP2002193996-A.	
XX	PD	10-JUL-2002.	
XX	PP	21-DEC-2000; 2000JP-00405097.	
XX	PR	21-DEC-2000; 2000JP-00405097.	
XX	PA	(KYOU) UNIV KYOTO.	
XX	DR	WPI: 2002-685438/74.	
XX	PT	P-PSDB; ABG71264.	
XX	PT	Glycinin, beta-conglycinin and proglycinin, their crystal structures, three dimensional coordinates, three dimensional structured and models	

PT and their uses.

PS Disclosure; Page 1264-1266; 1298pp; Japanese.

CC The present invention relates to a new Glycinin characterised by the
CC atomic coordinate data fully defined in the specification. The structure
CC can be used for improving processability of soya protein. The present
CC nucleic acid sequence encodes the Glycine max (Soybean) var.
CC Shiotsurunoko protein #1, as described in the specification

SQ Sequence 1488 BP; 466 A; 368 C; 362 G; 292 T; 0 U; 0 Other;

Query Match 33.7%; Score 513; DB 6; Length 1488;

Best Local Similarity 62.7%; Pred. NO. 2.6e-136;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;

QY 5 AGCAACCGAGGAGAACGGTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTGACAATC 64

Db 71 AGCAGCCTCAGCAAAACGAGTGCCAGATCCAAAACTCAATGCCCTCAAACCGGATAACC 130

QY 65 GCATTGAATCAGAGGGCGGTTACATTGAGACTTGGAACCCCAACACAGGAGTTCGAAT 124

131 GTATAGAGTCAGAAGGAGGGCTCATTTGAGACATGGAACCTTAACAACAGCCATTCCAGT 190

QY 125 GCGCGGCGTGGCCCTCTCTCGCTTAGTCTCGCGCGCAAGCCCTTCGTAGGCCCTTCT 184

Db 191 GTGCCGGTGTGCCCCCTCTCTCGCTGCACCCCTCAACCGCAACGCCCTTCGTAGACCTTCCT 250

Qy 185 ACTCAAATGCTCCCGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTTGATAT 244

Db 251 ACACCAACGGTCCCCAGGAAATCTACATCCAAGGTAAGGTAATTTTGGCATGATAT 310

QY 245 TCCCTGGTTGCTCCTAGACACTATGAAGAGCCTCACACACAGGTCGTCGATCTCAGTCCC 304

Db 311 ACCCGGGTTGTCTCCTAGCACATTGAAGAGCCTCA----- 344

QY 305 AAAGACCACCAAGACGCTCTCCAAGGAGAAGACCAAAGCCACAGCAACGAGATAGTCACC 364

Dbb 345 -----ACAACCTCAACAAGAGGACAAAGCAGCAGACCAAGACCGTCACC 391

QY 365 AGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCACCGGTGCTTTCT 424

Db 392 AGAAGATCTATAACTTCAGAGAGGGTGATTGTATCGCAGTGCCTACTGGTGTTCATGGT 451

QY 425 GGCTCTACAACGACCACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACA 484

Ddb 452 GGATGTACAACAATGAAGACACTCCTGTTGTTGCCGTTTCTATATTGACACCAACAGCT 511

QY 485 ACGACAAC CAGCTTGATCAGTTC CCGAGGAGATTCAATTTGGCTGGGAACACGGAGCAAG 544

D_b 512 TGGAGAACCCAGCTCGACCAGATGCCCTAGGAGATTCTATCTTGTGGGAACCAAGAGCAAG 571

QY 545 AGTTCTTAAGGTACCAAGCAACAAGCAGACGAAGCAAGCTTACCATATAGCC 604

Db 572 AGTTTCTAAAAATATCAGCAAGAG----- 594

Qy 605 CATAAGCCCGCAAGTCAGCCTAGACAAGAAGAGCGTGAAATTAGCCCTCGAGGACAGC 664

Db 595 -----CAAGGAGGTCATC 607

QY 665 ACAGCCG CAGAGAA C GAGCAGCAAGAGAAACGAGGTGGAACATCTTCAGCG 724

Db 608 AAAGCCAGAAAGGAAAGCATCAGCAAGAAAGAAACGAAGGAGGCAGCATATTGAGTG 667

QY 725 GCTTCACGCCGGAGTTCCTGGAACAAGCCTTCCAGGTTGACGACAGACAGATAGTGCAA 784

Db 668 GCTTCACCCCTGGAAATTCCTTGGAACATGCATTCAGCGT---GGACAAGCAGATAGCGAAA 724

QY 785 ACCTAAGAGCGAGACCGAGAGTGAAGAAGAGGGAGCCATTGTGACAGTGAAGGGAGGCC 844

Db 725 ACCTACAAGGAGAGAACGAAGGGGAAGACAAGGAGCCATTGTGACAGTGAAGGAGGTC 784

QY 845 TCAGAACTTTGAGCCCA-----GATAGAAAGAGACGTGCCGACGAAGAAGAGGAAT 895

XX	20-MAR-2002; 2002DE-01012892.	Db	251	ACACCAACGGTCCCCAGGAAATCTATACCAACAAGGTAAAGGTATTTTGGCATGATAT	310
XX	(BADI) BASF PLANT SCI GMBH.	Qy	245	TCCTGGTTGTCTAGACACTATGAAGAGCCTCACACACAAGGTGCTCGATCTCAGTCCC	304
XX		Db	311	ACCGGGTTGTCTTAGCACAATTTGAAGACCTCA	344
XX	Kock M, Bauer J;	Qy	305	AAAGACCAACCAAGACGTCTCCAAGGAGAAAGACCAAGCCAAACAGCAACGAGATAGTCACC	364
XX	WPI; 2003-803889/75.	Db	345	-----ACAACTCAACAAGAGGACAAAGCAGCAGACCAAGACCGTCACC	391
XX	P-PSDB; ADH89245.	Qy	365	AGAAGTGACACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCGCCCGGTGGCTTTCT	424
XX	Reducing expression of at least two target genes, useful e.g. for	Db	392	AGAAATCTATAAATCTCAGAGAGGGTGATTTGATCGCAGTGGCTACTGGTGTGTCATG	451
XX	producing transgenic plants, using partly double-stranded interfering	Qy	425	GGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACAACA	484
XX	RNA.	Db	452	GGATGTACAACAATGAAGACACTCTCTGTTGTTGCGGTTTCTATTATTGACACCAACAGCT	511
XX	Claim 7; SEQ ID NO 19; 228pp; German.	Qy	485	ACGACAAACAGCTTGTATCAGTTCCCGCAGGAGATTTCAATTTGGCTGGGAACAACGGAGCAAG	544
XX	This invention describes a novel method for reducing the expression of at	Db	512	TGGAGAACAGCTCGACAGATGCCTAGGAGATTCTATCTTGTCTGGAAACCAAGAGCAAG	571
XX	least two different endogenous target genes in a eukaryotic cell or	Qy	545	AGTTCTTAAGTTACAGCAACAAAGCAGACAAAGCAGACAAAGAGCTTACCATATAGCC	604
XX	organism by introducing an RNA molecule that is at least partly double	Db	572	AGTTTCTAAATAATATCAGCAAGAG-----	594
XX	stranded. The transcribed RNAs from at least two target genes have	Qy	605	CATACAGCCCGCAAAAGTCAAGCTCAGCCTAGACAAGAAGCGTGAAATTTAGCCCTCGAGGACAGC	664
XX	homology below 90% and the RNA molecule is formed as a single, self-	Db	595	-----	607
XX	complementary molecule. At least one of the double-stranded structures	Qy	665	ACAGCCGACAGAAACGAGCAGGACGAAGAAGAAAGAAACGAAGGTGGAACATCTTTCAGCG	724
XX	formed from individual sense sequences has an even number of repeats of	Db	608	AAAGCCAGAAAGGAAGACATCAGCAAGAAGAAGAAACGAAGAGGAGCAGCATATTGAGTG	667
XX	21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At	Qy	725	GCTTACGCCGAGTTCTCTGGAACAAGCCTTCAGGTTGACGACACAGATAGTCAAA	784
XX	least two target genes are selected from different classes of storage	Db	668	GCTTACCCCTGGAATTTCTTTGGAACATGCAATTCAGCGT---GGACAAGCAGATAGCGA	724
XX	protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-	Qy	785	ACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCGACATTTGTACAGTTGAGGGAGGCGC	844
XX	prolamine and at least one of the sense sequences is identical to storage	Db	725	ACCTACAGAGAGAAAGAAAGGGAAGACAAGGAGCCATTTGTACAGTTGAAGAGGAGTTC	784
XX	protein sequences or genes in the homogenitstate metabolic pathway or	Qy	845	TCAGATCTTTGAGCCCA-----GATAGAAAGAGACGTCCCGACGAAAGAGAGAAAT	895
XX	enzymes types, e.g. acetyl transacylases, chioesterases, (de)branching	Db	785	TGAGCGTGATAAAACCCACCACCGACGACGACCAACAAGAGCCCGAGGAAGAGAAG	844
XX	expression systems, vectors and transgenic organisms are used for	Qy	896	AGATGAACATGAATA-----TGAATACGATGAAGAGGATAGAGGCGTG	940
XX	preparation of pharmaceuticals, in biotechnological processes and plant	Db	845	AAAGAGAGGATGAGAAGCCACAGTGCAGGGTAAAGACAAACACTGCCAACCCCGGAG	904
XX	biotechnology, specifically in plants to improve protection against	Qy	941	GCAGGGGAAGCAGAGCAGCGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTA	1000
XX	abiotic stress, to modify composition and/or content of fatty acids,	Db	905	GAAGCCAAAGCAAAAGCAGAGAAATGGCAATTTGACGAGACCATATGCACCATGAGACTTC	964
XX	lipids and oils, to modify carbohydrate composition, to alter colour or	Qy	1001	AAAAGAACATTTGTTAGAAAACAGATCCCTTGACATCTTACAACCCCTCAAGTGTGTTACTCA	1060
XX	pigmentation, to reduce content of storage proteins, to increase	Db	965	GCCACAACATTTGGCCAGACTTTCATCACCTGACATCTACAACCTCAAGCCGGTAGGCTCA	1024
XX	resistance to pathogens, to inhibit stem break, to delay fruit ripening	Qy	1061	AAACTGCCAACCATCTCAACCTTCTTAATTAATTTAGTGGGCTTGGACCTAGTCTGATATG	1120
XX	or aging, to induce male sterility, to reduce content of toxic or	Db	1025	CAACGCCCAACGAGCTTGAATTTCCCAAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTTG	1084
XX	unwanted components, to modify lignification and/or lignin content, to	Qy	1121	GAATCTCTACAGGAATGCAATTTGTTGTCGCTCACTTACAACACCAACGACACAGCATCA	1180
XX	susceptibility to shock, to increase synthesis of Vitamin E, to reduce	Db	1085	GATCTCTCCGCAAGATGCAATGTTTCGTCCACACTACAACCTGAAACGCAACAGCATAA	1144
XX	contents of nicotine, caffeine or theophylline and to increase methionine	Qy	1181	TATATCGATTGAGGGAGCGGGCTCAGTGCAGAGTCTGTGACAGCAACGCGCAACAGAGTGT	1240
XX	content, by reducing threonine biosynthesis. The method provides a rapid	Db	1145	TATACGCAATTTGAATGGACGGGCATTGATACAAAGTGTGTGAATTCGAACCGGTGAGAGAGTGT	1204
XX	and efficient way of reducing gene expression, can inhibit more than one	Qy	1241	ACGACGAGAGCTTCAAGAGGGTCACTGCTGTTGTTGTCACAGACTTTCGCGCTCGCTG	1300
XX	target gene, prevents development of multiple phenotypes (since the	Db	1205	TTGATGGAGAGCTGCAAGAGGGACGGGTGCTGATCTGTCGCAAAAACCTTTTGGTGGCTG	1264
XX	transcription rate is the same for all RNA sequences, significantly				
XX	reducing the selection process required to produce an organism with				
XX	effective suppression of all target genes), avoids problems of epigenic				
XX	gene silencing, does not require synthesis of individual RNA sequences				
XX	and the method can be applied to plants with complex (polyploid) genomes.				
XX	No interference between the individual RNA sequences occur. This sequence				
XX	represents a target gene used in the method of the invention.				
SQ	Sequence 1488 BP; 466 A; 368 C; 362 G; 292 T; 0 U; 0 Other;				
	Query Match 33.7%; Score 513; DB 10; Length 1488;				
	Best Local Similarity 62.7%; pred. No. 2.6e-136;				
	Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;				
QY	5 AGCAACCGGAGGAGAACCGCTGCGCAGTTCCAGCGCTCAATTCGCGAGAGACCTTGACAATC	64			
Db	71 AGCAGCCTCAGCAAAACGAGTGCAGATCCAAAAAATCTCAATGCCCTCAAAACCGGATAACC	130			
QY	65 GCATTGAATCAGAGGGGGTTACATTGAGACTTGGAAACCCCAACAACACGAGGTTTCAAT	124			
Db	131 GTATAGAGTCAGAAAGAGGGCTCATTTGACATATGAAACCTTAACAACAGCAATTCAGT	190			
QY	125 GCGCCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTTCGTAGGCTTTCT	184			
Db	191 GTGCCGCTGTTGCCCTCTCTCGCTGACCTCAACCGACGCGCTTTCGTAGACCTTCTCT	250			
QY	185 ACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGTTGATAT	244			

[illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 17:25:50 ; Search time 3574.43 Seconds
(without alignments)
16229.158 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcacacaccgagagagaa.....ctccgaggctgtggtcttaa 1524

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651.6	42.8	770	6	CD038329
2	604.4	39.7	725	6	CD038327
3	562.2	36.9	676	6	CD038254
4	554.6	36.4	710	6	CD038107
5	548.6	36.0	718	6	CD038332
6	487	32.0	553	7	CD087496
7	436	28.6	602	6	CD038301
8	417.6	27.4	532	6	CD038222
9	386.2	25.3	720	6	CD038043
10	373.6	24.5	452	7	CD087497
11	372	24.4	679	6	CD038144
12	330.4	21.7	906	5	BQ123409
13	321.6	21.1	974	5	BQ123283
14	316.2	20.7	886	6	CD038043
15	314	20.6	861	6	CD038043
16	310.6	20.4	810	6	CD038379
17	310.4	20.4	862	6	CD038379
18	310.2	20.4	907	5	BQ123241
19	307.8	20.2	822	5	BQ123014
20	305	20.0	787	6	CD038379
21	305	20.0	812	5	BQ123297
22	304.4	20.0	782	6	CD038379
23	304.4	20.0	879	5	BQ123329
24	303.6	19.9	824	6	CD038379

ALIGNMENTS

RESULT 1
CD038329
LOCUS CD038329 770 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI007 B05 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI007_B05 5', mRNA sequence.
ACCESSION CD038329
VERSION CD038329.1 GI:30420167
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
REFERENCE 1 (bases 1 to 770)
AUTHORS Luo M., Dang P., Guo B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
CONTACT: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1..770
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI007 B05"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="XL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80oC freezer. Total RNA was isolated with TRIzol-Reagent and ultrapurified (GIBCOBRL). mRNA was extracted and purified from

FEATURES
source

CD038329 19..9 881 6 CA858490
25 302.6 19..9 881 6 CA858490
26 298.4 19..6 790 6 CA858419
27 296.4 19..4 836 6 CA858453
28 296.2 19..4 812 4 B1311450
29 294 19..3 805 6 CA857850
30 293.4 19..3 827 6 CA858430
31 292.6 19..2 879 6 CA858252
32 292.4 19..2 789 4 B1310542
33 291.6 19..1 769 5 BQ124073
34 291 19..1 903 6 CA858778
35 290.8 19..1 879 6 CA858395
36 288.8 19..0 892 6 CA858816
37 288.6 18..9 746 4 B1311303
38 288.6 18..9 787 4 B1311483
39 288.2 18..9 793 6 CA857838
40 287.6 18..9 541 6 CD038610
41 287.4 18..9 752 6 CA858702
42 287.2 18..8 755 6 CA858724
43 286.4 18..8 756 6 CA857897
44 286.2 18..8 811 5 BQ123350
45 285.2 18..7 796 6 CA857752

total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise phagescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match		42.8%;	Score 651.6;	DB 6;	Length 770;
Best Local Similarity		92.9%;	Pred. No. 1.7e-168;		
Matches 694;		Conservative 0;	Mismatches 44;	Indels 9;	Gaps 1;
QY	391	GATCTCATGTCAGTTCCACCGGTGTTGCTTTCGCGCTCTCAACGACGACGACCTGAT	450		
DB	33	GATCTCATNGNNNACCGGCTCGAGGAATTCGGCACGAGGCAACGACCATGACACTGAT	92		
QY	451	GTGTTGCTGCTTCTCTTACTGACACCAACAAACGACGACGAGCTTGATCAGTTCCCC	510		
DB	93	GTGTTGCTGCTTCTCTTACTGACACCAACAAACGACGACGAGCTTGATCAGTTCCCC	152		
QY	511	AGGAGATTCAATTTGGCTGGGAACACGGGACGAGAGTTCTTAAGGTACCAGCAACAAGC	570		
DB	153	AGGAGATTCAATTTGGCTGGGAACACGGGACGAGAGTTCTTAAGGTACCAGCA	205		
QY	571	AGCAAGACGACGAGAGAGCTTACATATAGCCATACAGCCGCAAGTTCAGCTTAGA	630		
DB	206	--ACAAAGCAGACGAAAGAGCTTACCATATAGCCATACAGCCGCAAGTTCAGCTTAGA	263		
QY	631	CAAGAGAGGCTCAATTTAGCCCTCAGGACGACACAGCCGCGAGAGAAAGAGGACAA	690		
DB	264	CAAGAGAGGCTCAATTTAGCCCTCAGGACGACACAGCCGCGAGAGAAAGAGGACAA	323		
QY	691	GAAGAAGAAAACGAAGGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCTTGGAAACA	750		
DB	324	GAAGAAGAAAACGAAGGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCTTGGCAACA	383		
QY	751	GCCTTCAGGTTGACACACAGATAGTGCAAAACCTAAGAGCGAGACCGAGAGTGAA	810		
DB	384	GCCTTCAGGTTGACACACAGATAGTGCAAAACCTAAGAGCGAGACCGAGAGTGAG	443		
QY	811	GAAGAGGAGCCATTGTGACAGTGAGGAGGCTCAGAAATCTTAAGAGCGAGACCGAGAG	870		
DB	444	GAAGAGGAGCCATTGTGACAGTGAGGAGGCTCAGAAATCTTAAGAGCGAGACCGAGAG	503		
QY	871	AGACGTGCCGACGAAGAAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGAT	930		
DB	504	AGAGGTGCCGACGAAGAAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGAT	563		
QY	931	AGNAGCGTGGCAGGGAAGCAGAGCGAGGGGAATGGTATTGAACAGACGATCTGCACC	990		
DB	564	AGAAGCGTGGCAGGGAAGCAGAGCGAGGGGAATGGTATTGAACAGACGATCTGCACA	623		
QY	991	GCAAGTGTAAAAGAACATTGGTAGAAACAGATCCCTGACATCTACAACTCAAGCT	1050		
DB	624	GCAAGTGTAAAAGAACATTGGTAGAAACAGATCCCTGACATCTACAACTCAAGCT	683		
QY	1051	GGTTCACTCAAAACTGCCAACGATCTCAACCTTCTAATATTAGTGGCTTGGACCTAGT	1110		
DB	684	GGTTCACTCAAAACTGCCAACGATCTCAACCTTCTAATATTAGTGGCTTGGACCTAGT	743		
QY	1111	GCTGAATATGGAATCTCTACAGGAT	1137		
DB	744	GCTGAATATGGAATCTCTACAGGAT	770		

RESULT 2

CD038327

LOCUS

CD038327 725 bp mRNA linear EST 07-MAY-2003

DEFINITION

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES

source

1. .725
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI008_F09"
/tissue_type="Immature pods"
/dev_stage="R6"
/lab_host="xL1-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -800C freezer. Total RNA was isolated with TRIZOL-Reagent ultrapure (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 39.7%; Score 604.4; DB 6; Length 725;
Best Local Similarity 92.8%; Pred. No. 1.9e-155;
Matches 648; Conservative 0; Mismatches 41; Indels 9; Gaps 1;

QY	2	GGCAGCAACCGGAGGAGAACCGGTGCCAGTTCCAGCGCTCAATCGCAGACCTTGACA	61
DB	37	GGCAGCAGCGGAGGAAATGCGTCCAGTTCAGCGCTCAATCGCAAGGCGCTGACA	96
QY	62	ATCGCATTAATCAGAGGCGGTACATTGAGACTTGAACCCCAACACGAGGAGTTGG	121
DB	97	ACCGCATTTGAATCGGAGGCGGTACATTGAGACTTGAACCCCAACACGAGGAGTTGG	156
QY	122	AATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCTTT	181
DB	157	AATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGGCTTT	216


```
|||||
Db      654 AGGGGACGGGCTCATGTGCAAGT 676

RESULT 4
CD038107
LOCUS
DEFINITION
  CD038107 710 bp mRNA linear EST 07-MAY-2003
  Arachis hypogaea cDNA clone UTPPI002_E08 5', mRNA sequence.
ACCESSION
  CD038107
VERSION
  CD038107.1 GI:30419945
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
  1 (bases 1 to 710)
  Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
  and Lynch, R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd, Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3.
  Location/Qualifiers
    1..710
    /organism="Arachis hypogaea"
    /mol_type="mRNA"
    /cultivar="A13"
    /db_xref="taxon:3818"
    /clone="UTPPI002_E08"
    /tissue_type="Immature pods"
    /dev_stage="R6"
    /lab_host="XL1-blue"
    /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
    (UTPP)"
    /notes="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
    cDNA library was constructed from peanut cultivar A13
    (NCV1XR4). A13 has resistance to Aspergillus infection
    and drought tolerance. The immature pods that developed to
    R6 stage were collected from different plants, and placed
    into liquid N2 immediately and stored in -80oC freezer.
    Total RNA was isolated with TRIzol-Reagent
    ultrapure(GIBCOBRL). mRNA was extracted and purified from
    total RNA (Promega). cDNA synthesis and library
    construction followed the protocol of by ZAP-cDNA Gigapack
    III Gold cloning kit (Stratagene). The cDNA above 500bp
    were collected after size-fraction. The inserts were
    directionally cloned into Uni-ZAP XR vector using XhoI
    EcoRI sites adapters. The lambda library was packed into
    phages using Gigapack III Gold (Stratagene). The
    un-amplified library was used to excise phagescript
    phagemids from the Uni-ZAP XR vector, and the phagemids
    was used to transform the host bacteria SOLR. The library
    was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN
Query Match 36.4%; Score 554.6; DB 6; Length 710;
Best Local Similarity 98.3%; Pred. No. 1.1e-141;
Matches 571; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 945 GGGAGCAGAGCGGGAATCGTATTGAAGAGCATCTGCACCGCAAGTGCTAAAAA 1004
Db 1 GGGAGCAGAGCGGGAATCGTATTGAAGAGCATCTGCACCGCAAGTGTTAAAAA 60
QY 1005 GAACATTGGTAGAAACAGATCCCTTGACATCTACAACCCCTCAAGCTGTTCACATAAAC 1064
|||||
Db      61 GAACATTGGTAGAAACAGATCCCTTGACATCTACAATCTCAAGCTGTTCACATAAAC 120
QY 1065 TGGCAACGATCTCAACCTTCTTAATATTAGTGGCTTGGACCTAGTGTGTAATATGGAAA 1124
Db 121 TGGCAACGATCTCAACCTTCTTAATCTTATAGTGGCTTGGACTTAGTGTGAATATGGAAA 180
QY 1125 TCTCTACAGGAATGCAATGTTTGTTCGCTCAGTACCAACCAACGACACAGCATCATATA 1184
Db 181 TCTCTACAGGAATGCAATGTTTGTTCGCTCAGTACCAACCAACGACACAGCATCATATA 240
QY 1185 TCGATTGAGGGGACGGGCTCAGTGAAGTCTGTGACAGCAACGGCAACAGAGTGTACGA 1244
Db 241 TCGATTGAGGGGACGGGCTCAGTGAAGTCTGTGACAGCAACGGCAACAGAGTGTACGA 300
QY 1245 CGAGGAGCTTCAAGAGGGTTCAGTCTTGTGTGTCACAG-AAC TTCGCCGCTCGCTGGAA 1303
Db 301 CGAGGAGCTTCAAGAGGGTTCAGTCTTGTGTGTCACAGNAAC TTCGCCGCTCGCTGGAA 360
QY 1304 AGTCCAGAGCGAGAACTTCGAATACGTGGCATTCGAAGACAGACTCAAGGCCAGCATAG 1363
Db 361 AGTCCAGAGCGACAACTTCGAATACGTGGCATTCGAAGACAGACTCAAGGCCAGCATAG 420
QY 1364 CCAACTCGCGGTGAAAACTCCGTATAGATAA CTTGCCGAGGAGGTGTGCAAAATT 1423
Db 421 CCAACTCGCGGTGAAAACTCCGTATAGATAA CTTGCCGAGGAGGTGTGCAAAATT 480
QY 1424 CATATGGCTCCAAAGGGAGCGAGGCAAGCAGCTTAAGAACAAACCCCTTCAAGTTCT 1483
Db 481 CATATGGCTCCCAAGGGAGCGAGGCAAGCAGCTTAAGAACAAACCCCTTCAAGTTCT 540
QY 1484 TCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 541 TCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 581

RESULT 5
CD038332
LOCUS
DEFINITION
  CD038332 718 bp mRNA linear EST 07-MAY-2003
  UTPPI004_B04 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
  Arachis hypogaea cDNA clone UTPPI004_B04 5', mRNA sequence.
ACCESSION
  CD038332
VERSION
  CD038332.1 GI:30420170
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
ORGANISM
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
  1 (bases 1 to 718)
  Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
  and Lynch, R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3.
  Location/Qualifiers
    1..718
    /organism="Arachis hypogaea"
    /mol_type="mRNA"
    /cultivar="A13"
    /db_xref="taxon:3818"
    /clone="UTPPI004_B04"
    /tissue_type="Immature pods"
    /dev_stage="R6"
    /lab_host="XL1-blue"

FEATURES
source
  1..718
  /organism="Arachis hypogaea"
  /mol_type="mRNA"
  /cultivar="A13"
  /db_xref="taxon:3818"
  /clone="UTPPI004_B04"
  /tissue_type="Immature pods"
  /dev_stage="R6"
  /lab_host="XL1-blue"
```

/clone.lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/notes="vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV11X4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapur (GIBCOBRL). mRNA was extracted and purified from total RNA (Promega). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA GigaPack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using GigaPack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 36.0%; Score 548.6; DB 6; Length 718;
Best Local Similarity 93.4%; Pred. No. 4.9e-140;
Matches 621; Conservative 0; Mismatches 29; Indels 15; Gaps 4;

QY 875 GTCCCGACGAAGAGAGGATACGATGAAGATGAATATGAATACGATGAAGAGAGGAGGC 927
DB 1 GTCCCGACGAAGAGAGGATACGATGAAGATGAATATGAATATGAATGAAGAGAGGAGGC 60

QY 928 -----GATAGAGCGTGGCAGGGGAGCAGAGCGGGGGAATGGTATTGAAGAGACGA 982
DB 61 AACAAAGATAGAAGCGTGGCAGGGGAGCAGAGCGGCAATGGCAATGGAGGAGACCA 120

QY 983 TCTGCACCGCAGTGTCTAAAGAAACATCTGTAGAACACATCCCTGACATCTACAACC 1042
DB 121 TCTGCACCGCAGTGTCTAAAGAAACATCTGTAGAACACATCCCTGACATCTACAACC 180

QY 1043 CTCAGCTGGTTCACTCAAAACCTGCCAACGATCTCAACCTTTCTAATCTTAGTGGCTTG 1102
DB 181 CTCAGCTGGTTCACTCAAAACCTGCCAACGATCTCAACCTTTCTAATCTTAGTGGCTTG 240

QY 1103 GACCTAGTCTGAATATGGAATCTCTACAGGAATGCAATGTTTGGCTCACTACACAA 1162
DB 241 GACCTAGTCTGAATATGGAATCTCTACAGGAATGCAATGTTTGGCTCACTACACAA 300

QY 1163 CCAACGCACACACATCATATATCGATTGAGGGGACGGGCTCACGTGCAAGTGTGACAA 1222
DB 301 CGAACGCACACACATCATATATCGATTGAGGGGACGGGCTCACGTGCAAGTGTGACAA 360

QY 1223 GCAACGSCAACAGAGTGTACGAGGAGGCTTCAAGAGGGTCAAGTGTGTTGGTGGCAC 1282
DB 361 GCAACGSCAACAGAGTGTGTGAGGAGGCTTCAAGAGGGTCAAGTGTGTTGGTGGCAC 420

QY 1283 AG-AACTTCCCGTGGTGAAGTCCAGAGCGAGAACTTCAAGTACGTGGCAATCAAG 1341
DB 421 AGAAACTTCCCGTGGTGAAGTCCAGAGCGAGAACTTGAATACTGGCAATCAAG 480

QY 1342 ACAGACTCAAGGGCCAGCATAGCAACTCTCGCGGTGAAACTCCCGTCATAGATAACCTG 1401
DB 481 ACAGACTCAAGGGCCAGCATAGCAACTCTCGCGGTGAAACTCTCTTCATAGATAACTG 540

QY 1402 CC-GGAGGAGGTGGTGAAGATTCATATGGCCCTCCAAAGGGAGCAGGCAAGGAGCTTAA 1460
DB 541 CCGGAGGAGGTGGTGAAGATTCATATGGCCCTCCAAAGGGAGCAGGCAAGGAGCTTAA 600

QY 1461 GAACAACAA-CCCTTCAAGTCTTCTGTTCCACCGTCTCAGAGTCTCCGAGGGCTGTGG 1519
DB 601 GAACAACAAACCCCTTCAAGTCTTCTGTTCCACCGTCTCAGAGTCTCTCAGGGCTGTGG 660

QY 1520 CTTAA 1524
|||||

Db 661 CTTAA 665

RESULT 6
CO897496LOCUS
DEFINITION

CO897496 553 bp mRNA linear EST 13-AUG-2004
EST00001 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.

ACCESSION
VERSION

CO897496
CO897496.1 GI:51237286

KEYWORDS

EST.

SOURCE

Arachis hypogaea (peanut)

ORGANISM

Arachis hypogaea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.

REFERENCE

1 (bases 1 to 553)

AUTHORS

Yan, Y.S., Wang, L. and Huang, S.Z.

TITLE

Expressed sequence tags of mid-matured cotyledons of peanut

JOURNAL

Unpublished (2004)

COMMENT

Contact: Shangzhi Huang

COMMENT

Plant Development and Molecular Biology

COMMENT

Sun Yat-sen University, Department of Biology and Biotechnology

COMMENT

Guangzhou, 510275, China

COMMENT

Tel: (860) 02084036592

COMMENT

Email: YYS9803@yahoo.com.cn

COMMENT

Seq primer: pTriplex2 Forward.

FEATURES

Location/Qualifiers

source

1..553

source

/organism="Arachis hypogaea"

source

/mol_type="mRNA"

source

/cultivar="Yueyou 523"

source

/db_xref="taxon:3818"

source

/cissue_type="Cotyledons"

source

/dev_stage="Mid-matured stage"

source

/lab_host="E.coli BM25.8"

source

/clone_lib="Peanut Lambda Express library"

source

/note="Organ: Seed; Vector: lambdaTriplex2"

ORIGIN

Query Match 32.0%; Score 487; DB 7; Length 553;
Best Local Similarity 93.6%; Pred. No. 4.8e-123;
Matches 508; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 39 CCTCAATGCGCAGAGACCTGACAATCGCATTCGAATCAGAGGGCGGTTACATTGAGACTTG 98
DB 1 CCTCAATGCGCAGAGACCTGACAATCGCATTCGAATCAGAGGGCGGTTACATTGAGACTTG 60

QY 99 GAACCCCAACAACAGAGGTTGGAATGCGCGGCGCTCTCTCGCTTAGTCTCTCG 158
DB 61 GAACCCCAACAACAGAGGTTGGAATGCGCGGCGCTCTCTCGCTTAGTCTCTCG 120

QY 159 CGCAACGCGCTTCGTTAGGCTTTCTACTCCAAATGCTCCCAAGAGATCTTTCACGCA 218
DB 121 CGCAACGCGCTTCGTTAGGCTTTCTACTCCAAATGCTCCCAAGAGATCTTTCACGCA 180

QY 219 AGGAAGGGGATATCTTTGGGTTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCTCA 278
DB 181 AGGAAGGGGATATCTTTGGGTTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCTCA 240

QY 279 CACACAAGGTCTCGATCTCAGTCCCAAGAGACCAAGAGCTCTCAAGGAGAGACCA 338
DB 241 ACAACAAGAGGCGCGCATCAGTCCCAAGAGACCAAGAGCGTTTCAAGGAGAGACCA 300

QY 339 AAGCCAAACAGCAACAGATAGTACCAGAGGTGACCGTTTCGATGAGGGTATCTCAT 398
DB 301 AAGCCAAACAGCAACAGATAGTACCAGAGGTGACCGTTTCGATGAGGGTATCTCAT 360

QY 399 TGCAGTTCCACCGGTGTTGCTTTCTGGCTCTAACAGCAGCAGCACTGATGTTGTGC 458
DB 361 TGCAGTTCCACCGGTGTTGCTTTCTGGCTCTAACAGCAGCAGCAGCAGTATGTTGTGC 420

```

QY 459 TGTTCCTCTTACTGACACCAACAAACGACCAACCCAGCTTGATCAGTTCCCCAGGAGATT 518
Db 421 TGTTCCTCTTACTGACACCAACAAACGACCAACCCAGCTTGATCAGTTCCCCAGGAGATT 480

QY 519 CAATTTGGCTGGGAACACGAGGACGAAGTCTTAAAGTACCAGCAACAAAGCGACAAAG 578
Db 481 CAATTTGGCTGGGAACACGAGGACGAAGTCTTAAAGTACCAGCAACAAAGCGACGAAG 540

QY 579 CAG 581
Db 541 AAG 543

RESULT 7
CD038301
LOCUS
DEFINITION
  CD038301 602 bp mRNA linear EST 07-MAY-2003
  Arachis hypogaea cDNA clone UTPPI005_C08 5', mRNA sequence.
ACCESSION
  CD038301
VERSION
  CD038301.1 GI:30420139
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
REFERENCE
  1 (bases 1 to 602)
  Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
  and Lynch,R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3
  Location/Qualifiers
    1..602
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPPI005_C08"
      /tissue_type="Immature pods"
      /dev_stage="R6"
      /lab_host="XLI-blue"
      /clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library
      (UTPP)"
      /notes="Vector: Uni-ZAP XR; Site_1: EcoRI; Site_2: XhoI;
      cDNA library was constructed from peanut cultivar A13
      (NCVixAR4). A13 has resistance to Aspergillus infection
      and drought tolerance. The immature pods that developed to
      R6 stage were collected from different plants, and placed
      into liquid N2 immediately and stored in -80oC freezer.
      Total RNA was isolated with TRIzol-Reagent
      ultrapure(GIBCOBRL). mRNA was extracted and purified from
      total RNA (Promega). cDNA synthesis and library
      construction followed the protocol of by ZAP-cDNA Gigapack
      III Gold cloning kit (Stratagene). The cDNA above 500bp
      were collected after size-fraction. The inserts were
      directionally cloned into Uni-ZAP XR vector using XhoI
      EcoRI sites adapters. The lambda library was packed into
      phages using Gigapack III Gold (Stratagene). The
      un-amplified library was used to excise phlambdaescript
      phagemids from the Uni-ZAP XR vector, and the phagemids
      was used to transform the host bacteria SOLR. The library
      was constructed by Dr. Meng Luo and Dr. Phat Dang."
FEATURES
  source
    1..602
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPPI005_C08"
      /tissue_type="Immature pods"
      /dev_stage="R6"

```

```

Query Match 28.6%; Score 436; DB 6; Length 602;
Best Local Similarity 95.7%; Pred. No. 6.2e-109;
Matches 448; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 1057 CTCAAAACTGCCAACGATCTCAACCTTCTAATACCTTAGTGGCTTAGTGTCTGAA 1116
Db 1 CTCAAAACTGCCAACGAGCTCAACCTTCTAATACCTTAGTGGCTTAGTGTCTGAA 60

QY 1117 TATGGAATCTCTACAGGAATGATTTGTTGCTCACTACAACACCAACACACACAGC 1176
Db 61 TATGGAATCTCTACAGGAATGATTTGTTGCTCACTACAACACGACACACAGC 120

QY 1177 ATCATATATCGAATGAGGGACGGGCTCACGTGCAAGTCGTGCAGACAGCAACGGCAACAGA 1236
Db 121 ATCATATATCGAATGAGGGACGGGCTCACGTGCAAGTCGTGCAGACAGCAACGGCAACAGA 180

QY 1237 GTGTACGACGAGGAGCTTCAAGAGGGTCACGTGCTTGTGTGTCGCCACAGAACTTCGCCGTC 1296
Db 181 GTGTTTCGACGAGGAGCTTCAAGAGGGTCACGTGCTTGTGTGTCGCCACAGAACTTCGCCGTC 240

QY 1297 GCTGGAAGTCCACAGAGCGAGAACTTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCC 1356
Db 241 GCTGGAAGTCCACAGAGCGAGAACTTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCC 300

QY 1357 AGCATAGCCAACTCGCCGGTGAAAACTCCGTCATAGATAAATCTCCGAGGAGGTGTT 1416
Db 301 AGCATAGCCAACTAGCCGGTGAAAACTCCGTCATAGATAAATCTCCGAGGAGGTGTT 360

QY 1417 GCAATTCATATGGCTTCCAAAGGAGCGAGGCAAGGAGCTTAAGAACACAAACCCCTTC 1476
Db 361 GCAATTCATATGGCTTCCAAAGGAGCGAGGCAAGGAGCTTAAGAACACAAACCCCTTC 420

QY 1477 AAGTTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 421 AAGTTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 468

RESULT 8
CD038222
LOCUS
DEFINITION
  CD038222 532 bp mRNA linear EST 07-MAY-2003
  Arachis hypogaea cDNA clone UTPPI004_B11 5', mRNA sequence.
ACCESSION
  CD038222
VERSION
  CD038222.1 GI:30420060
KEYWORDS
  EST.
SOURCE
  Arachis hypogaea (peanut)
  Arachis hypogaea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
  Aeschynomeneae; Arachis.
REFERENCE
  1 (bases 1 to 532)
  Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
  and Lynch,R.E.
  Generation and Analyses of ESTs for Arachis hypogaea
  Unpublished (2003)
  Contact: Baozhu Guo
  Molecular Genetics
  USDA/ARS, Crop Protection and Management Research Unit
  2747 Davis Rd., Tifton, GA 31794, USA
  Tel: 229-387-2334
  Fax: 229-387-2321
  Email: bguo@tifton.usda.gov
  Seq primer: T3
  Location/Qualifiers
    1..532
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPPI004_B11"
      /tissue_type="Immature pods"
      /dev_stage="R6"
FEATURES
  source
    1..532
      /organism="Arachis hypogaea"
      /mol_type="mRNA"
      /cultivar="A13"
      /db_xref="taxon:3818"
      /clone="UTPPI004_B11"
      /tissue_type="Immature pods"
      /dev_stage="R6"

```

/lab host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapur (GIBCOBRL). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 27.4%; Score 417.6; DB 6; Length 532;
Best Local Similarity 96.8%; Pred. No. 7.3e-104; Indels 0; Gaps 0;
Matches 426; Conservative 0; Mismatches 14;

QY 1085 TAATCTAGTGGCTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCATTGT 1144
DB 1 TAATCTAGTGGCTGGACTAGTCTGAATATGGAATCTCTACAGGAATGCATTGT 60
QY 1145 TTGTGCTCACTCAACCAACGACACAGCATATATATGATGAGGGACGGGCTC 1204
DB 61 TTGTGCTCACTCAACCAACGACACAGCATATATATGATGAGGGACGGGCTC 120
QY 1205 ACCTGCAAGTCGTGGACAGCAACGCAAGAGTGACGAGGAGCTTCAAGAGGCTC 1264
DB 121 ACCTGCAAGTCGTGGACAGCAACGCAAGAGTGACGAGGAGCTTCAAGAGGCTC 180
QY 1265 ACCTGCTTGTGGTCCACAGAACTTCGCCGTGGAAAGTCCAGAGCGAGAACTTCG 1324
DB 181 ACCTGCTTGTGGTCCACAGAACTTCGCCGTGGAAAGTCCAGAGCGAGAACTTCG 240
QY 1325 AATACGTGGCATTCAGACAGACTCAAGGCCAGCAGTAGCAACCTCGCCGGTGAAACT 1384
DB 241 AATACGTGGCATTCAGACAGACTCAAGGCCAGCAGTAGCAACCTCGCCGGTGAAACT 300
QY 1385 CCGTCATAGTAACCTCGCGAGGAGGTGCTGCAATTCATATGCTCCAAAGGAGC 1444
DB 301 CCGTCATAGTAACCTCGCGAGGAGGTGCTGCAATTCATATGCTCCAAAGGAGC 360
QY 1445 AGCAAGGAGCTTAAGCAACAAACCCCTTCAAGTCTTCGTTCCACCGTCTCAGCAGT 1504
DB 361 AGCAAGGAGCTTAAGCAACAAACCCCTTCAAGTCTTCGTTCCACCGTCTCAGCAGT 420
QY 1505 CTCGAGGGCTGTGGCTTAA 1524
DB 421 CTCGAGGGCTGTGGCTTGA 440

RESULT 9
CD038043
LOCUS CD038043 720 bp mRNA linear EST 07-MAY-2003
DEFINITION UTPPI001_F09 USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI001_F09 5', mRNA sequence.
ACCESSION CD038043
VERSION CD038043.1 GI:30419881
KEYWORDS EST.
SOURCE Arachis hypogaea (peanut)
ORGANISM Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

rosids; eurousids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 720)
Luo,M., Dang,P., Guo,B.Z., Holbrook,C.C., Lee,R.D., Bausher,M.G.
and Lynch,R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov

FEATURES

source

Location/Qualifiers
1..720
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPP1001_F09"
/tissue type="Immature pods"
/dev stage="R6"
/lab_host="XLI-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI; cDNA library was constructed from peanut cultivar A13 (NCV1XAR4). A13 has resistance to Aspergillus infection and drought tolerance. The immature pods that developed to R6 stage were collected from different plants, and placed into liquid N2 immediately and stored in -80°C freezer. Total RNA was isolated with TRIzol-Reagent ultrapur (GIBCOBRL). cDNA synthesis and library construction followed the protocol of by ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The cDNA above 500bp were collected after size-fraction. The inserts were directionally cloned into Uni-ZAP XR vector using XhoI EcoRI sites adapters. The lambda library was packed into phages using Gigapack III Gold (Stratagene). The un-amplified library was used to excise pBluescript phagemids from the Uni-ZAP XR vector, and the phagemids was used to transform the host bacteria SOLR. The library was constructed by Dr. Meng Luo and Dr. Phat Dang."

ORIGIN

Query Match 25.3%; Score 386.2; DB 6; Length 720;
Best Local Similarity 96.8%; Pred. No. 3.8e-95;
Matches 394; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1118 ATGGAATCTCTACAGGAATGCATTGTTGTGCTCCTCACTACACACCAACGACACAGCA 1177
DB 1 ATGGAATCTCTACAGGAATGCATTGTTGTGCTCCTCACTACACACCAACGACACAGCA 60
QY 1178 TCATATATCGATTGAGGGACGGGCTCAGCTGCAAGTCGTGGACAGCAACGGCAACAGAG 1237
DB 61 TCATATATCGATTGAGGGACGGGCTCAGCTGCAAGTCGTGGACAGCAACGGCAACAGAG 120
QY 1238 TGTACGACGAGGAGCTTCAAGAGGGTCAGCTGCTTGTGTGCTGCCACAGAACTTCGCCGTG 1297
DB 121 TGTACGACGAGGAGCTTCAAGAGGGTCAGCTGCTTGTGTGCTGCCACAGAACTTCGCCGTG 180
QY 1298 CTGGAAGTCCACAGACGAGAACTTCGAATACGTGTCATTCAAGACAGACTCAAGGCCCA 1357
DB 181 CTGGAAGTCCACAGACGAGAACTTCGAATACGTGTCATTCAAGACAGACTCAAGGCCCA 240
QY 1358 GCATAGCCAACTCGCCGGTGAAACTCCGTTCATAGATAACCTGCCGGAGGAGGTGTTG 1417
DB 241 GCATAGCCAACTCGCCGGTGAAACTCCGTTCATAGATAACCTGCCGGAGGAGGTGTTG 300
QY 1418 CAAATTTCATATGCGCTCCAAAGGAGCAGGACGCTTAAGACACACACCCCTTCA 1477

301 CAAATTATATGCTCCCAAGGAGCAGGAGGAGCTTAAAGAACAAACCCCTTCA 360
1478 AGTTCCTCGTCCACCGTCTCACAGTCTCCGAGGCTGGGCTTAA 1524
361 AGTTCCTCGTCCACCTTTTTCAGAGTCTCCGAGGCTGGGCTTAA 407

RESULT 10
CO897497

LOCUS
DEFINITION
CO897497
VERSION
KEYWORDS
SOURCE
ORGANISM

CO897497 452 bp mRNA linear EST 13-AUG-2004
EST00002 Peanut Lambda Express library Arachis hypogaea cDNA 5',
mRNA sequence.
CO897497.1 GI:51237287
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 452)
Yan, Y.S., Wang, L. and Huang, S.Z.
Expressed sequence tags of mid-matured cotyledons of peanut
Unpublished (2004)
Contact: Shangzhi Huang
Plant Development and Molecular Biology
Sun Yat-sen University, Department of Biology and Biotechnology
Guangzhou, 510275, China
Tel: (860) 02084036592
Email: yys9803@yahoo.com.cn
Seq primer: pTriplEx2 Forward.
Location/Qualifiers
1. .452
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="Yueyou 523"
/db_xref="taxon:3818"
/issue_type="Cotyledons"
/dev_stage="Mid-matured stage"
/lab_host="E.coli BM25.8"
/clone_lib="Peanut Lambda Express library"
/notes="Organ: Seed; Vector: lambdaTriplEx2"

FEATURES
source

ORIGIN

Query Match 24.5%; Score 373.6; DB 7; Length 452;
Best Local Similarity 91.7%; Pred. No. 1e-91;
Matches 407; Conservative 0; Mismatches 34; Indels 3; Gaps 1;

QY 2 GGCAGCAACCGAGGAGAACGCGTGCAGTTCAGCGCCTCAATGCGCAGAGACCTGACA 61
DB 9 GGCAGCAGCGGAGGAGATGCGTGCGATTCAGCGCCTCAATGCGCAGAGCCTGACA 68
QY 62 ATCGCAATTGAATCAGAGGCGGTTAATTGAGACTTTGGAACCCCAACACCGAGGTTCC 121
DB 69 ACCGCTTGAATCGGAGGCGGTTACATTGAGACTTTGGAACCCCAACACCGAGGTTCC 128
QY 122 AATGCGCGGCGTGGCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCCTT 181
DB 129 AATGCGCGGCGTGGCCCTCTCGCGTTAGTCTCGCGCGCAACGCCCTTCGGAGCCCTT 188
QY 182 TCTACTCCCAATGTCTCCAGGAGATCTTCAATCAGCAAGGAGGATACTTTGGGTGCA 241
DB 189 TCTACTCCCAATGTCTCCAGGAGATCTTCAATCAGCAAGGAGGATACTTTGGGTGCA 248
QY 242 TATTCCCTGGTCTCTAGACACTATGAGAGCCTCACACAAAGTCTGTCGATCTCAGT 301
DB 249 TATTCCCTGGTCTCTAGACACTATGAGAGCCTCACACAAAGTCTGTCGCGCATCACT 308
QY 302 CCAGAGACCAACCAAGCTCTCCAGGAGAGAGACCAAGCCCAACAGCAACGA 358
DB 309 CGAAAGAGACCAACCAAGCTTTTGAAGGAGAGAGACCAAGCCCAACAGCAACGA 368

QY 359 GTCCACCAAGGTGCACCGTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTG 418
DB 369 GTCCACCAAGAGTGCCTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTG 428
QY 419 CTTTCTGGCTTACAAACGACCAGC 442
DB 429 CTCTCTGGATGTACAACGACCATG 452

RESULT 11
CD038144
LOCUS
DEFINITION
SOURCE
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CD038144 679 bp mRNA linear EST 07-MAY-2003
UTPPI003 All USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)
Arachis hypogaea cDNA clone UTPPI003_All 5', mRNA sequence.
CD038144
CD038144.1 GI:30419982
EST.
Arachis hypogaea (peanut)
Arachis hypogaea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae;
Aeschynomeneae; Arachis.
1 (bases 1 to 679)
Luo, M., Dang, P., Guo, B.Z., Holbrook, C.C., Lee, R.D., Bausher, M.G.
and Lynch, R.E.
Generation and Analyses of ESTs for Arachis hypogaea
Unpublished (2003)
Contact: Baozhu Guo
Molecular Genetics
USDA/ARS, Crop Protection and Management Research Unit
2747 Davis Rd., Tifton, GA 31794, USA
Tel: 229-387-2334
Fax: 229-387-2321
Email: bguo@tifton.usda.gov
Seq primer: T3.
Location/Qualifiers
1. .679
/organism="Arachis hypogaea"
/mol_type="mRNA"
/cultivar="A13"
/db_xref="taxon:3818"
/clone="UTPPI003_All"
/issue_type="Immature pods"
/dev_stage="R6"
/lab_host="Xll-blue"
/clone_lib="USDA-Tifton Peanut Immature Pod cDNA Library (UTPP)"
/note="Vector: Uni-ZAP XR; Site 1: EcoRI; Site 2: XhoI;
cDNA library was constructed from peanut cultivar A13
(NCV11XAR4). A13 has resistance to Aspergillus infection
and drought tolerance. The immature pods that developed to
R6 stage were collected from different plants, and placed
into liquid N2 immediately and stored in -80oC freezer.
Total RNA was isolated with TRIZOL-Reagent
ultrapure (GIBCOBRL). mRNA was extracted and purified from
total RNA (Promega). cDNA synthesis and library
construction followed the protocol of by ZAP-cDNA Gigapack
III Gold cloning kit (Stratagene). The cDNA above 500bp
were collected after size-fraction. The inserts were
directionally cloned into Uni-ZAP XR vector using XhoI
EcoRI sites adapters. The lambda library was packed into
phages using Gigapack III Gold (Stratagene). The
un-amplified library was used to excise Bluescript
phagemids from the Uni-ZAP XR vector, and the phagemids
was used to transform the host bacteria SOLR. The library
was constructed by Dr. Meng Luo and Dr. Phat Dang."

FEATURES
source

ORIGIN

Query Match 24.4%; Score 372; DB 6; Length 679;
Best Local Similarity 92.9%; Pred. No. 3.2e-91;
Matches 390; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1105 CCTAGTCTGATGTAATGAAATCTCTACAGGAATGCATTTGTTGCGTCTACTACAACACC 1164
 |||||
 Db 46 CCTGAGGCTGCAGGAATTCGCGACGAGGGGAATGCATTTGTTGTCCTCTACTACAACACG 105
 |||||
 Qy 1165 AACGCACACAGCATCATATATCGATTGAGGGGCGGCTCACGTCGCAAGTCGTGGACAGC 1224
 |||||
 Db 106 AACGCACACAGCATCATATATGCAATTGAGGGGCGGCTCATGTGCAAGTCGTGGACAGC 165
 |||||
 Qy 1225 AACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTCGTTGTGTGGTCCACAG 1284
 |||||
 Db 166 AACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTCGTTGTGTGGTCCACAG 225
 |||||
 Qy 1285 AACTTCGCGTCTGCGTGAAGTCCAGAGCGAGAACTTCGATACGTCGGCATTCAGACACA 1344
 |||||
 Db 226 AACTTCGCGTCTGCGTGAAGTCCAGAGCGAGAACTTCGATACGTCGGCATTCAGACACA 285
 |||||
 Qy 1345 GACTCAAGGGCCAGCATAGCAACCTCGCGGTGAAATCCGTCATAGATAAACCCTGCCG 1404
 |||||
 Db 286 GACTCAAGGGCCAGCATAGCAACCTCGCGGTGAAATCCCTCATAGATAAACCCTGCCG 345
 |||||
 Qy 1405 GAGAGGTGTTGCAAAATTCATATGGCTCTCAAGGAGGAGGCAAGGCGCTTAAGAAC 1464
 |||||
 Db 346 GAGAGGTGTTGCAAAATTCATATGGCTCTCAAGGAGGAGGCAAGGCGCTTAAGAAC 405
 |||||
 Qy 1465 AACAAACCCCTCAAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 1524
 |||||
 Db 406 AACAAACCCCTCAAGTCTTCTGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCTTAA 465
 |||||

RESULT 12

BQ123409
 LOCUS BQ123409
 DEFINITION EST608985 GLSD Medicago truncatula cDNA clone pGLSD-32121, mRNA sequence.

ACCESSION BQ123409

VERSION BQ123409.1

KEYWORDS GI:20175371

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE

1 (bases 1 to 906)

Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T.,

Cheung, F. and Fraser, C.M.

ESTs from late stage developing seeds of Medicago truncatula

Unpublished (2002)

Contact: Grusak, M.A.

USDA/ARS Children's Nutrition Research Center

Baylor College of Medicine

1100 Bates Street, Houston, TX 77030-2600, USA

Tel: 713 798 7044

Fax: 713 798 7078

Email: mgrusak@bcm.tmc.edu

TIGR sequence name: MTRAQ59TK More information is available at:

www.medicago.org

Seq primer: SKmd (CTA GAA CTA gtg gAT CC).

FEATURES

source
 1..906
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pGLSD-32121"
 /tissue_type="Immature seeds"
 /dev_stage="25 to 35 days after pollination"
 /lab_host="XLOLR"
 /clone_lib="GLSD"
 /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; Immature seeds, collected from pods, ranging in age
 from 25 to 35 days after pollination, were harvested from

greenhouse-grown plants. Seed were removed and separated
 from pod walls and immediately frozen in liquid nitrogen.
 Seeds throughout the age range were pooled for mRNA
 extraction. cDNA was prepared from polyA+ enriched RNA.
 The cDNA was directionally ligated into the Unizap XR
 vector from Stratagene and packaged using Gigapack III
 Gold packaging extracts. Plasmids containing cDNA inserts
 were excised from the recombinant lambda-Zap phage using
 Ex-assist helper phage and propagated in XL0LR cells."

ORIGIN

Query Match 21.7%; Score 330.4; DB 5; Length 906;
 Best Local Similarity 63.5%; Pred. No. 1.1e-79;
 Matches 540; Conservative 0; Mismatches 301; Indels 9; Gaps 2;
 Qy 680 GAGCAGGACAAGAGAAAGAAACGAAAGGTGGAAACATCTTTCAGCGGCTTTCAGCGCGAGT 739
 |||||
 Db 47 GAGGAAGGGAACAACAAGAAATATGAAGGAGGCAACATTTTCAGTGGCTTCAAGAGGGATT 106
 |||||
 Qy 740 TCCTGGACNAGCCTTCCAGGTTGACGACAGACAGATAGTGCAAAACCTTAAGAGGCGAGA 799
 |||||
 Db 107 TCTTGAAGATGCATTTGAACGTGA---ACAGGCATATAGTTGAAAACCTTCAAGGCGAGGA 163
 |||||
 Qy 800 CCGAGAGTGAAGAGAGGAGGAGCCATTGTGACGTGAGGGGAGCGCTCAGAAATCTTGAGCC 859
 |||||
 Db 164 ATGAAGACAGGAGAGAGGAGCCATTGTCAAAGTGAAGAGGAGACTCAGCATCATGAGCC 223
 |||||
 Qy 860 C-----AGATAGAAAGAGACGTGCCGACGAAAGAGGAATACGATGAAGATGAATATG 913
 |||||
 Db 224 CTCGAGAGAGACAACAACGCCACCCAGCAGACAGATGAAGATGAAGATGAAGAGG 283
 |||||
 Qy 914 AATAAGATGAAGAGGATAGAAAGCGGTGGCAGGGGAAAGCAGAGCAGGGGGAATGGTATTG 973
 |||||
 Db 284 ATGATGGAGGCCACATCATCAGAAAGAGAGAGAAAGAGAAAGGACACACAGCGTCTTG 343
 |||||
 Qy 974 AAGAGACGATCTGCACGCGCAAGTCTAAAGAAACATTTGTTAGAAACAGATCCCTGCACA 1033
 |||||
 Db 344 CGGAACAATTTGCACCTGCCAGGCTTCCAGAACATGGTCTCATCTTCACTACCTGACA 403
 |||||
 Qy 1034 TCTACAACCTCAAGCTGGTTCTACTCAAAACCTGCCAACGATCTCAACCTTCTTAATCTTA 1093
 |||||
 Db 404 TCTACAACCTCAAGCTGGTTAGAAATCAAAACATATCAACAGCTTCGACCTCCAGCTCTCA 463
 |||||
 Qy 1094 GGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCATTTGTTGTGCTC 1153
 |||||
 Db 464 GGTGGCTTAAGACTCAGTGTCTGAGCATGGAACTCTCCACAGAAATGCTATGTTGTTCCAC 523
 |||||
 Qy 1154 ACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTCACGTGCAAG 1213
 |||||
 Db 524 ACTACAATCTCAATGCAAAACAGCATATATTTGGCATTTGAATGGACGTGCAAGGCTACAG 583
 |||||
 Qy 1214 TCGTGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAGTGTGTTG 1273
 |||||
 Db 584 TAGTGAATGCAATGCGAACACACTGTGTTCGATGAGAACTAGAAAGCTGGCGCGTGCATTGA 643
 |||||
 Qy 1274 TGGTGCCACAGAACTTCGCGCTCGCTGGAAAGTCCGACAGCGAGAACTTCGAATACGTGG 1333
 |||||
 Db 644 TCGTGCCACAAAACCTTTGCTGTTGACGCAAAATCAGTGAAGCGACAGGTTCACTTATGTTT 703
 |||||
 Qy 1334 CATTCAAGACAGACTCAAGGCCGACGATAGCAACCTCGCGGTGAAACCTCCGTCATAG 1393
 |||||
 Db 704 CATTCAAGACCAATGATTAATGCGCAATTTGCCAGGCTTGCGAGGACACAATCCACTCTAA 763
 |||||
 Qy 1394 ATAACTGCGGAGGAGGTGGTTGCAAAATTCATATGGCTTCCAAAGGAGGACAGCAAGGC 1453
 |||||
 Db 764 GTGGTATGCCAGTGGATGTGCTTGCAGCTACATTTCAACATGGACAGGAATGAGGCAAGGC 823
 |||||
 Qy 1454 AGCTTAAGAACAAACACCCCTTCAAGTTCCTTCGTTCCACGCTCTCAGCAGTCTCCGAGGG 1513
 |||||
 Db 824 AGCTCAAAAACCAACATCTCTTTTAATTTCTAGTTCCACCCCGTGAGTCCGAAACGACAG 883
 |||||
 Qy 1514 CTGTGGCTTA 1523
 |||||

Db	884	CTCAGCTTA 893	
RESULT 13			
BQ123283			
LOCUS		974 bp mRNA linear EST 17-APR-2002	
DEFINITION		EST608859 GLSD Medicago truncatula cDNA clone pGLSD-31N1, mRNA	
ACCESSION		BQ123283	
VERSION		BQ123283.1	GI:20175245
KEYWORDS		EST:	
SOURCE		Medicago truncatula (barrel medic)	
ORGANISM		Medicago truncatula	
REFERENCE			
AUTHORS		Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., and Fraser, C.M.	
TITLE		ESTs from late stage developing seeds of Medicago truncatula	
JOURNAL		Unpublished (2002)	
COMMENT		Contact: Grusak, M.A. USDA/ARS Children's Nutrition Research Center Baylor College of Medicine 1100 Bates Street, Houston, TX 77030-2600, USA Tel: 713 798 7044 Fax: 713 798 7078 Email: mgrusak@bcm.tmc.edu TIGR sequence name: MTRAO73TK More information is available at: www.medicago.org Seq primer: SKmod (CTA GAA CTA gtg gat CC).	
FEATURES		Location/Qualifiers	
source		1..974	
		/organism="Medicago truncatula"	
		/mol_type="mRNA"	
		/cultivar="A17"	
		/db_xref="taxon:3880"	
		/clone="pGLSD-31N1"	
		/tissue_type="Immature seeds"	
		/dev_stage="25 to 35 days after pollination"	
		/lab_host="XLOLR"	
		/clone_lib="GLSD"	
		/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."	
ORIGIN			
Query Match		21.1%; Score 321.6; DB 5; Length 974;	
Best Local Similarity		63.7%; Pred. NO. 3e-77;	
Matches 524; Conservative		0; Mismatches 289; Indels 9; Gaps 2;	
QY	680	GAGCAGCACAAGAAAGAAACGAAGGTGGAAACATCTTTCAGCGGCTTCAGCCGGAGT	739
Db	102	GAGGAGGGNACACAAAGAAATGAGAGGGACACATTTTCAGTGGCTTCAGAGGGATT	161
QY	740	TCTTGGAACAAGCTTCCAGTTGACGACAGACAGATAGTGCAAAACCTTAAGAGGCGACA	799
Db	162	TCTTGAAGATGCAATCAACGT---GAACAGGCATATAGTTGAAAACCTTCAAGGCAGGA	218
QY	800	CCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGGAGCCCTCAGACTTTCGAGCC	859
Db	219	ATGAACACCAGGAAGAGGGAGCCATTGTCAAGTGGAGAGGACTCAGCATCATGAGCC	278
QY	860	C-----AGATAGAAAGAGACGTCGCGACGAAGAAGAGGAATACGATGAAGATGAATATG	913
Db	279	CTCCAGAGAGACAACAACGCCCCAGCAGACACAAGATGAAGAAGATGAAGATGAAGAGG	338
QY	914	AATACGATCAAGAGGATAGAAGCGTGGCAGGGGAGCAGAGCGAGGGGAATGGTATTG	973
Db	339	ATCAGTGGAGGCCACATCATCAGAAAGCAGAGAGAAAGAAAGGCAACACGGCTTTG	398
QY	974	AAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACATTCGTAGAAACACAGATCCCTTGACA	1033
Db	399	CGGAAACAATTTGCACTGCCAGGCTTCACGAGACATGCTCTCATCTTCATCACCTTGACA	458
QY	1034	TCTACAACCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAACCTTCTTAATATCTTA	1093
Db	459	TCTACAACCTCAAGCTGTTAGAAATCAAAACTATCACAGCTTCGACCTCCCGAGCTCTCA	518
QY	1094	GGTGGCTTGAGACTAGTGTGTAATATGGAATCTCTACAGGAATGATTTGTTGTGCTC	1153
Db	519	GGTGGCTAAGACTCAGTGTGAGCATGGAACCTCTCCACAGAAATGCTATGTTGTTCCAC	578
QY	1154	ACTACAACACCAACGCCACACACATCATATATCGATTGAGGGGAGGGCTCAGCTGCAAG	1213
Db	579	ACTACAATCTCAATGCAACAGCAATAATTTGGCATTTGAATGGACGTGCAAGGCTACAAG	638
QY	1214	TGCTGCACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTTCAGTGTG	1273
Db	639	TAGTGAACCTGCAATGGCAACACTGTGTTGATGAGGAACCTAGAGCTGGCCGTGCATTGA	698
QY	1274	TGCTGCACAGCAACCTTGGCCGTGGTGGAAAGTCCAGAGCGAGAACTTTCGGAATAGCTGG	1333
Db	699	TGCTGCACCAAAACTTTGCTGTTGCAGCAAAAATCAGTGAGCGACAGGTTTCACTTATGTTT	758
QY	1334	CATTCAAGACACACTCAAGGCCACGATAGCAACCTCCCGGTGAAACTTCGTGCATAG	1393
Db	759	CATTCAAGACCAATGATATGCCGCAATTCGACAGGCTTCAGGGACACAACTCACTCTTAA	818
QY	1394	ATAACCTCGCGGAGGAGTGGTTGCAAAATTCATATGGCTCCAAAGGGGAGCAGGCAAGGC	1453
Db	819	GTGGTATGCCAGTGGATGTGTTGCAGCTTACATTCAACTGACAGAGGAATGAGGCGAGC	878
QY	1454	AGCTTAAGAACCAACCCCTTCAAGTTCTTCTGTTCCACCGT	1495
Db	879	AGCTCAAAAACCAACATCTCTTTAATTTCTAGTTCAACCCCGT	920
RESULT 14			
CA857975		886 bp mRNA linear EST 18-DEC-2002	
EST635230		GLSD Medicago truncatula cDNA clone pGLSD-26J23, mRNA	
LOCUS			
DEFINITION		sequence.	
ACCESSION		CA857975	
VERSION		CA857975.1	GI:27232339
KEYWORDS		Medicago truncatula (barrel medic)	
SOURCE		Medicago truncatula	
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.	
REFERENCE		1 (bases 1 to 886)	
AUTHORS		Grusak, M.A., Samac, D., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., and Fraser, C.M.	
TITLE		ESTs from late stage developing seeds of Medicago truncatula	
JOURNAL		Unpublished (2002)	
COMMENT		Contact: Grusak, M.A. USDA/ARS Children's Nutrition Research Center Baylor College of Medicine 1100 Bates Street, Houston, TX 77030-2600, USA Tel: 713 798 7044 Fax: 713 798 7078 Email: mgrusak@bcm.tmc.edu TIGR sequence name: MTRM60TK More information is available at: www.medicago.org	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 23, 2005, 20:00:36 ; Search time 250.731 Seconds
(without alignments)
9945.676 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 cgcagcaaccggaggagaa.....ctccgaggcgtgtggtttaa 1524

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	517.4	34.0	1446	3	US-09-108-0108-13
2	517.4	34.0	1446	4	US-09-758-652-13
3	517.4	34.0	1446	4	US-10-684-651-13
4	513	33.7	1488	3	US-09-108-0108-11
5	513	33.7	1488	4	US-09-758-652-11
6	513	33.7	1488	4	US-10-684-651-11
7	513	33.7	1743	4	US-09-762-381-1
8	504	33.1	1458	3	US-09-108-0108-12
9	504	33.1	1458	4	US-09-758-652-12
10	504	33.1	1458	4	US-10-684-651-12
11	198	13.0	1551	3	US-09-108-0108-15
12	198	13.0	1551	4	US-09-758-652-15
13	198	13.0	1551	4	US-10-684-651-15
14	153.2	10.1	1689	3	US-09-108-0108-14
15	153.2	10.1	1689	4	US-09-758-652-14
16	153.2	10.1	1689	4	US-10-684-651-14
17	142	9.3	1706	4	US-09-462-720-1
18	136.6	9.0	1685	1	US-08-486-721A-1
19	136.2	8.7	1556	1	US-08-486-721A-2
20	114.6	7.5	3113	1	US-08-146-422-20
21	114.6	7.5	3113	1	US-08-626-554-2
22	108.4	7.1	4939	4	US-09-645-593-8
23	93.8	6.2	7218	1	US-08-232-463-14
24	53.4	3.5	3477	4	US-09-462-720-3
25	49.8	3.3	15252	4	US-09-949-016-13584
26	49.4	3.2	150394	4	US-09-949-016-13042
27	47.2	3.1	289	3	US-09-007-005-17

Sequence 17, Appl
Sequence 14193, A
Sequence 13840, A
Sequence 1, Appl
Sequence 33, Appl
Sequence 13032, A
Sequence 14, Appl
Sequence 439, App
Sequence 2, Appl
Sequence 2, Appl
Sequence 14543, A
Sequence 12557, A
Sequence 2, Appl
Sequence 1, Appl
Sequence 11814, A
Sequence 17474, A
Sequence 17475, A

ALIGNMENTS

RESULT 1
US-09-108-0108-13
; Sequence 13, Application US/09108010B
; Patent No. 6362399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION NUMBER: US/09/108,010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-108-0108-13

Query Match 34.0%; Score 517.4; DB 3; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.1e-144;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

```
QY 5 AGCAACCGGAGGAGAACCGCTGCGAGTTCCAGGCGCTCAATGCGCAGAGACCTGCAATC 64
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 AGCAGCCACAGCAAAAGAGTCCAGATCCAAAGCTCAATGCCCTAAACCGGATAACC 130
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 65 GCATTGAATCAGAGGGGGTTACATTGAGACTTTGGAAACCCCAACACACAGGAGTTGCAAT 124
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 GTATAGAGTCAGAGGTGGCTTCAATTGAGACATGGAACCCCTAACAAAGCCATTCCAGT 190
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 GCCCGCGCTGCCCTCTCTCGCTAGTCTCCGCGCAACGCCCTTCGAGGCCCTTCT 184
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 GTCCGGTGTGCCCTCTCTCGCTGACCCCTCAACCGCAACGCCCTTCGAGACCTTCCT 250
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 185 ACTCCAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGTTGATAT 244
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 ACACCAACGCTCCCAAGGAGATCTACATCCAAAGGTAGTGATTTTGGCATGATAT 310
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 245 TCCTCGTTGTCTAGACACTATGAAGAGCTCACAACAGTCTCGATCTCAGTCCC 304
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 TCCCGGTTGTCTTAGCACATTTGAAGAGCTT----- 342
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 305 AAAGACCACCAAGAGCTCTCCAGGAGGAAGACCAAGCCAAACAGCAACGAGATAGTCACC 364
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 343 -----CAACAAAGGAAGAGCAGAGGCCCAAGACCGTCACC 382
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 365 AGAAGGTGCAACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCCAACCGGTGTTGCTTCT 424
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTGCAGTGGCAACCGGTTTGGCATACT 442
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 425 GGCTCTACAAACGACGACGACACTGTATGTTGTTGCTGTTTCTCTTACTGACACCAACAA 484
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 443 GGATGTACAAACATGAAGACACTCTCTGTTGTTGTCGGTCTTCTCTTATTGACACCAACAGCT 502
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 485 ACCAACACAGCTTGATCAGTCTCCAGGAGATTCAATTTGGCTGGGAACAACGAGCAAG 544
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 503 TCAGAAACCAAGCTCGAACAGATGCCCTAGGAGATTCTATCTTGCTGGGAACCAAGAGCAAG 562
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 545 AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGCAGCAAGCTTACCATATAGCC 604
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 563 AGTTCTACAGTATCAGCCACAGCAGCAGCAAG----- 595
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 CATACAGCCGAAAGTCAGCTAGCCTAGACAAAGAGAGCGTGAAATTAGCCCTCGAGGACAGC 664
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 596 -----GAGGTACTC 604
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 665 ACAGCCGAGAGAACGAGCAGGACAAGAAAGAAAAAGAGTGAACATCTTCAGCG 724
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 605 AAAGCCAGAAAGGAAGCGTCACCAAGAGAGAGAAACGAGGAGCAGCATATTGAGTG 664
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 725 GCTTCAACGCGGAGTTCTTGGAACAAGCCCTTCAGGTTGACGACAGACAGATGTCNAA 784
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 665 GCTTCGCCCGGAATCTTGGAACATGCGTTC---GTCGTGGACAGCAGATAGTCAGAA 721
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 785 ACCTAAGAGCGAGACGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGGAGGCC 844
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 722 AGCTAAGAGGTGAGAACGAAGAGGAAGAGAGAGGTTGCCATTGTGACAGTGAAGAGAGTGC 781
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 845 TCAGAACTTTGAGCCAGATAGAAAGAGAGAGCTGCCGACGAAGAGAGGAATACGATGAAG 904
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 782 TCAGCGTGATAGCCCAACCCACGGAAGAGCAGCAACAAAGACCCGAGGAGAGGAAGC 841
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 905 ATGAATATGAATACGATGAAGAGGATAGAAGGCGTGGCAGGGGAAGCAGAGGAGGGGA 964
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 842 CAGATTGTGACGAGAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 965 ATGGTATTGAGAGAGCATCTGACCGCAAGTGCTAAAAAGAACATTTGGTAGAAACAGAT 1024
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 887 ATGGCATTTGACGAGACCAATTGACAAATGAGACTTCGCCACAAACATTTGCCAGACTTCAT 946
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1025 CCCCTGACATCTCAACCCCTCAAGCTGGTTTCACTCAAAAACCTGCCAACGATCTCAACCTTC 1084
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 947 CACTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCGCTACAGCCTCGACTTCC 1006
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 1085 TAATACTTAGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGAAATGCAATTGT 1144
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1007 CAGCCCTCTCGTGGCTCAAACCTCAGTGGCCAGTTTTGGATCACTCCGCAAGAATGCTATGT 1066
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1145 TTGTGCTCACTACAAACACCAACGACACAGCATCATATATCGATTGAGGGACGGGCTC 1204
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1067 TCGTGCCACACTACAAACCTGAAACGCAACAGACATAATATACGCATTTGAATGGAACGGGCAT 1126
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1205 AGTGTCAAGTCTGTGACAGCAGCAACGAGAGTGTACGACGAGGAGCTTCAAGAGGGTTC 1264
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1127 TGGTACAGTGTGATTTGCAATGTTGAGAGAGTGTGTTGATGAGAGCTGCAAGAGGGAC 1186
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1265 AGTGTCTTGTGTGCCACAGAACTTCGCGTTCGCTGGAAAGTCCAGAGCGAGAACTTCG 1324
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1187 AGGTGTTAATTGTGCCACAAAACCTTTGCGTGGCTGCAAGATCACAGAGCGCAAACTTCG 1246
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1325 AATACTGTGCATTCAAGACAGACTCAAGCCCAAGATAGCAACCTCCGCGGTGAAACT 1384
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1247 AGTATGTTTCAATCAAGACCAATGATAGACCTCGATCGGCAACCTTCAGGTGCAAACT 1306
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1385 CCGTCAATAGATAACCTGCGGAGGAGTGGTTGCAAAATTCATATGGCCTCCAAAGGGAGC 1444
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1307 CATTGTTGACGCATTTGCGGAGGAGTGATTTCAGCAAACTTTTAACCTAAGAGGCGAGC 1366
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1445 AGGCAAGCAGCTTAAGAACAAACCCCTTCAAGTTCCTTCAAGTTCCTTCAAGTTCAGCAGT 1504
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1367 AGGCCAGGCGAGTCAAGAACAAACCCCTTTCAGCTTCTCGTTCCACCTCAAGAGGAGTCTC 1426
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1505 CTCGAGGGCTGTGCTTA 1523
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1427 AGAGGAGAGTGTGCTTA 1445
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-09-758-652-13
; Sequence 13, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
;
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; ;
```

		; TYPE: nucleic acid			
		; STRANDEDNESS: single			
		; TOPOLOGY: linear			
		; MOLECULE TYPE: CDNA			
		; SEQUENCE DESCRIPTION: SEQ ID NO: 13:			
		US-09-758-652-13			
		Query Match		34.0%; Score 517.4; DB 4; Length 1446;	
		Best Local Similarity		62.8%; Pred. No. 1.1e-144;	
		Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;			
QY	5	AGCAACCGGAGAGAACGGTTCAGTTCACGGCTCAATTCGCGCAGAGACCTGACAATC	64		
DB	71	AGCAGCCACAGCAAAACGAGTGCAGATCCAAACGCTCAATGCCCTAAACCGGATACC	130		
QY	65	GCATTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACCAACAGAGGTTGGAAT	124		
DB	131	GTATAGAGTCAGAAGGTGGCTTCATTGAGACATGGAAACCCCTAAACCAAGCCATTCCAGT	190		
QY	125	GCGCGGCGTGGCCCTCTCTCGTTAGTCTCTCGCGCGCAAGCCCTTCGTAGGCTTTCT	184		
DB	191	GTGCGGCGTGTGCGCTCTCTCGCTGCAACCCCTCAACCGCAACGCGCTTCGAGACCTTCCT	250		
QY	185	ACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGGAGGGATACCTTCGGTTGATAT	244		
DB	251	ACACCAAGCTGCCCAAGGAGATCTACATCCAAAGGTAGTGGTATTTTGGCATGATAT	310		
QY	245	TCCTCGTTGTCTTAGACACTATGAAGAGCCCTCACACAAAGGTCTCGATCTCAGTCCC	304		
DB	311	TCGCGGTTGTCTTAGACACTTTGAAGAGCT-----	342		
QY	305	AAAGACCACCAAGAGCTCTCAAGGAGAGACCAAAAGCCAAACAGCAGATAGTCACC	364		
DB	343	-----CAACAAAGGAGCAAGCAGCAGCGCCCAAGACCGTCAAC	382		
QY	365	AGAAGTGACCGTTTCGATGAGGTGATCTCATTTGCGATTTCCACCGGTGTTCTTCT	424		
DB	383	AGAAGATCTATCACTTTCAGAGAGGGTGAATTTGATTCGAGTGCACACCGTTTTGCATACT	442		
QY	425	GGCTCTACAACAGCACACACTGATGTTGTTGCTTCTTCTTACTGACACCAACAACA	484		
DB	443	GGATGTACAACATGAGACACTCTGTTGTCGCTTTCTTTATTGACACCAACAGCT	502		
QY	485	ACGACAAACGAGTTGATCAGTTCCTCCAGGAGATTCATTTGCGTGGGAACCGGAGCAAG	544		
DB	503	TCAGAAACGAGCTCGACACAGATGCTTAGAGATTTCTATCTTGTGGGAACCAAGAGCAAG	562		
QY	545	AGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCC	604		
DB	563	AGTTTCTACAGTATCAGCCACAGACAGCAGCAAG-----	595		
QY	605	CATACAGCCGCAAGTTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACAGC	664		
DB	596	-----GAGGTACTC	604		
QY	665	ACAGCCGACAGAACGAGCAGCAAGAGAGAAAGAAACGAAAGGTGGAAACATCTTCAGCG	724		
DB	605	AAAGCCAGAAAGGAAAGCGTCAGCAAGAAAGAAAGAAACGAAAGGAGGACGATATTGAGTG	664		
QY	725	GCTTCAGCGCGAGTTCTTGGAAACAGCCCTTCAGGTTGACGACGACAGATAGTGCAAA	784		
DB	665	GCTTCGCGCGGAAATTTTGGAAACATGCGTTC---GTCGTGACAGGACAGATAGTGAA	721		
QY	785	ACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGAGGCC	844		
DB	722	AGCTACAGGTGAGAACGAAAGAGGAGAGAGGGTGCATTGTGACAGTGAAGAGGTC	781		
QY	845	TCAGATCTTTGAGCCAGATAGAAAGAGACGCTGCCGACGAAAGAGAGGAATACATGAAG	904		
DB	782	TCAGCGTGAAGCCACCCACGGAAGAGCAGCAACAAAGACCCGAGGAGAGGAGAGC	841		
QY	905	ATGAATATGAATACGATGAAGAGGATAGAAAGCGCTGGCAGGGGGAAGCAGAGCGGGGA	964		

Db	842	CAGATTGTGACGAGAAAGACAAACAT-----TGCAAAAGCCAAAGACGAA	888
QY	965	ATGGTATTGAAGAGACGATCTGACCGCAAGTGCTAAAGAAAGCAATTTGGTAGAACACAGAT	1024
DB	887	ATGGCATTGACGAGACCATTTTGCACAATGAGACTTTCGCCCAACCATTTGGCCAGACATTCAT	945
QY	1025	CCCTTGACATCTACAACCTCAGCTGTTCACTCAAAACCTGCCAACGATCTCAACCTTC	1084
DB	947	CACCTGACATCTTCAACCTCAGCTGTTAGCATCACAAACGCTACCAAGCTTCGACTTCC	1006
QY	1085	TAATACCTTAGTGGCTTGACCTAGTCTGAAATATGGAATCTCTACAGGAATGCATTGT	1144
DB	1007	CAGCCCTCTGCTGGCTCAAACTCAGTGCCTGCTTGGATCACTCGCAGAAGATGCTATGT	1066
QY	1145	TTGTCGCTCACTACAACACCAACGACACAGCATCATATATCGATTGAGGGGACGGGCTC	1204
DB	1067	TCGTGCCACACTACAACCTGAAGCGCAACAGCATATATACGCAATTGAATGGAACGGGCAT	1126
QY	1205	AGTGCAGTCTGACAGCAACGCGCAACAGAGTGTACGACGAGAGCTTCAAGAGGCTC	1264
DB	1127	TGGTACAAGTGGTGAATTCGAATGGTGAAGAGTGTTTGATGAGAGAGCTGCAAGAGGAC	1186
QY	1265	AGTGTCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAGCTGCCAGAGCGAGAACTTCG	1324
DB	1187	AGTGTGAATTTGTGCCCAAAACTTTGCGGTGGCTGCAAGATCACAGACGCAACTTCG	1246
QY	1325	AATAGCTGCTTCAAGACAGACTCAAGCGCCAGCATAGCCAACTCGCCGTTGAAAACCT	1384
DB	1247	AGTATGTTTCACTTCAAGACCAATGATAGACCTCGATCGGCAACCTTGCAGTGCMAACT	1306
QY	1385	CGTCATAGATAACTCTGCGGAGAGGTGTTGCAAAATTCATATGCGCTCCAAAGGGAGC	1444
DB	1307	CATTGTTGAACGCAATTGCGGAGGAAGTGAATTCAGCAAACTTTTAACTAAGGAGGAGC	1366
QY	1445	AGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAGCAGT	1504
DB	1367	AGGCGAGGAGGTCAAGAACCAACACCCCTTTCAGCTTCTGTTCCACCTAAGGAGTCTC	1426
QY	1505	CTCCGAGGCGTGTGGCTTA 1523	
DB	1427	AGAGGAGAGTTGTGGCTTA 1445	
RESULT 3			
US-10-684-651-13			
; Sequence 13, Application US/10684651			
; Patent No. 6828491			
; GENERAL INFORMATION:			
APPLICANT: ANTHONY JOHN KINNEY			
GARY MICHAEL FADER			
TITLE OF INVENTION: SUPPRESSION OF SPECIFIC			
CLASSES OF SOYBEAN SEED			
PROTEIN GENES			
NUMBER OF SEQUENCES: 21			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY			
STREET: 1007 MARKET STREET			
CITY: WILMINGTON			
STATE: DELAWARE			
COUNTRY: UNITED STATES OF AMERICA			
ZIP: 19898			
COMPUTER READABLE FORM:			
MEDIUM TYPE: DISKETTE, 3.50 INCH			
COMPUTER: IBM PC COMPATIBLE			
OPERATING SYSTEM: MICROSOFT WINDOWS 95			
SOFTWARE: MICROSOFT WORD 7.0A			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/10/684,651			
FILING DATE: 14-Oct-2003			
CLASSIFICATION: <Unknown>			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US/09/758,652			
FILING DATE: 11-Jan-2001			

RESULT 3

US-10-684-651-13

; Sequence 13, Application US/10684651

; Patent No. 6828491

; GENERAL INFORMATION:

; APPLICANT: ANTHONY JOHN KINNEY

; APPLICANT: GARY MICHAEL FADER

; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC

; CLASSES OF SOYBEAN SEED

; PROTEIN GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY

; STREET: 1007 MARKET STREET

; CITY: WILMINGTON

; STATE: DELAWARE

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 19898

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.50 INCH

; OPERATING SYSTEM: MICROSOFT WINDOWS 95

; SOFTWARE: MICROSOFT WORD 7.0A

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/684,651

; FILING DATE: 14-Oct-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/758,652

; FILING DATE: 11-Jan-2001

COMPUTER READABLE FORM:	
MEDIUM TYPE: DISKETTE, 3.50 INCH	
COMPILER: IBM PC COMPATIBLE	
OPERATING SYSTEM: MICROSOFT WINDOWS 95	
SOFTWARE: MICROSOFT WORD 7.0A	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/09/108,010B	
FILING DATE: 30-Jun-1998	
CLASSIFICATION: <Unknown>	
PRIOR APPLICATION DATA:	
APPLICATION NUMBER: 60/019,940	
FILING DATE: JUNE 14,1996	
ATTORNEY/AGENT INFORMATION:	
NAME: LYNNE M. CHRISTENBURY	
REGISTRATION NUMBER: 30,971	
REFERENCE/DOCKET NUMBER: BB-1071-A	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: 302-992-5481	
TELEFAX: 302-773-0164	
INFORMATION FOR SEQ ID NO: 11:	
SEQUENCE CHARACTERISTICS:	
LENGTH: 1488 base pairs	
TYPE: nucleic acid	
STRANDEDNESS: single	
TOPOLOGY: linear	
MOLECULE TYPE: CDNA	
SEQUENCE DESCRIPTION: SEQ ID NO: 11:	
US-09-108-010B-11	
Query Match	33.7%; Score 513; DB 3; Length 1488;
Best Local Similarity	62.7%; Pred. No. 2.3e-143;
Matches 968; Conservative	0; Mismatches 425; Indels 150; Gaps 5;
QY	5 AGCAACGGAGAGAACGGTGCAGTTCACGGCTCAATGCGCAGAGACTGCACAATC 64
DB	71 AGCAGCTCAGCAACAGAGTCCAGATCCAAACTCATGCTCCAAACCGGATACC 130
QY	65 GCATTGAATCAGAGGCGGTTCATTGAGACTTGAACCCCAACCAACAGGAGTTCGAAT 124
DB	131 GTATAGAGTCAGAGGAGGGCTCATTTAGACATGGAACCCCTAAACAAGCCATTCAGT 190
QY	125 GCGCGGGTGGCTCTCTCGCTAGTCTCGCGGGCAAGCCCTTCGTAGGCTTTCT 184
DB	191 GTGCGGGTGTGCGCTCTCTCGCTGCAACCCCTCAACCGCAAGCCCTTCGTAGACCTTCT 250
QY	185 ACTCCAAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGGATCTTTGGGTGATAT 244
DB	251 ACACCAACGGTCCCAAGGAATCTCATCCAAAGAGTAAAGGTATTTTGGCAATGAT 310
QY	245 TCCTGGTGTGCTTACAGACTATGAGAGCTTCACACCAAGAGTTCGTGATCTCAGTCCC 304
DB	311 ACCCGGTTGCTTACAGCAATTTGAAGGCTCA----- 344
QY	305 AAGACCAACAGAGCTTCCAGGAGAGAACCAAGCCCAACAGCAGATAGTCAAC 364
DB	345 -----ACAACCTCAACAAAGAGGACAAAGCAGCAGACCAAGACCGTCAAC 391
QY	365 AGAAGGTGACCGTTTCGATGAGGTGATCTCATTTGAGTTCACACCGGTGCTTTCT 424
DB	392 AGAAGATCTTAATCTTACAGAGGGGTGATTTGATGCGAGTGCCTACTGGTGTGATGGT 451
QY	425 GGCTCTACACCAACCAACGACACTGATGTTGTTGCTTTCTTTACTGACACCAACACA 484
DB	452 GGATGTACAAACATGAAGACACTCCTGTTGTTGCGGTTCTATTATTGACACCAACAGCT 511
QY	485 ACAGAACACAGCTTGATGATGTTCCCGCAGGAGATTCATTTGCTGGGAAACCGAGCAAG 544
DB	512 TGGAGAACACAGCTTCGACCAAGATGCTTAGAGATTCATCTTGTGGGAAACCAAGCAAG 571
QY	545 AGTTCTTAAAGTACCAAGCAACAAAGCAGACCAAGCAGCAAGCAAGGAGTTCACCATATAGCC 604
DB	572 AGTTCTTAAATATACGCAAGAG----- 594

RESULT 6
US-10-684-651-11
; Sequence 11, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-684-651-11

Query Match 33.7%; Score 513; DB 4; Length 1488;
Best Local Similarity 62.7%; Pred. No. 2.3e-143;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;
QY 5 AGCAACCGGAGGAGAACCGCTGCCAGTTCAGCGCCTCAATGGCGAGAGACCTGACAAATC 64
DB 71 AGCAGCCTCAGCAAAACGAGTGGCAGATCCAAAACCTCAATGGCCTCAAAACCGGATAACC 130
QY 65 GCATTGAATCAGAGGGCGGTACATTGAGATTTGGACCTTGGACCCCAACACAGGAGTTCGAAT 124
DB 131 GTATAGAGTCAGAGGAGGGCTCATTTGAGACATGGAACCCCTCAACAAAGCCATTCCAGT 190
QY 125 GCGCGCGCTGCGCCCTCTCTCGTTAGTTCCTCGCCCGCAACGCCCTTCGTAGGCGCTTTCT 184
DB 191 GTGCGGTGTGCGCCCTCTCTCGTGCACCTCAACCGCAGCGCCTTCGTAGACCTTCCT 250
QY 185 ACTCCAAATGCTCCCGAGAGATCTTCATCCAGCAAGGAAGGAGATCTTTGGGTGATAT 244
DB 251 ACACCAACGGTCCCGAGGAATCTACATCCAAAGGAAGGTATTTTGGCATGATAT 310
QY 245 TCCTGCTGCTCTAGACATATGAGAGCCTCACACAAAGGTGCTGATCTCAGTCCC 304

DB 311 ACCGGGTTGCTCTAGCACATTTGAAGAGCCTCA----- 344
QY 305 AAAGACCAACAGAGCGTCTCCAAAGAGAGAACCAAGCCAAACAGCAACGAGATAGTCACC 364
DB 345 -----ACAACTCAAAAGAGGACAAAGAGCAGACAGACCAAGACCGTCACC 391
QY 365 AGRAGGTGCACCGTTTTCGATGAGGGTGATCTCATTCAGCTTCCCAACGGTGTGCTTCT 424
DB 392 AGAAGATCTATAACTTCAGAGAGGGTGATTTGATCGAGTGCTCTACTGGTGTGCTAGGT 451
QY 425 GGCTCTACAAACGACACGACACTGATGTTGTTCTGCTGTTTCTTCTTACTGACACCAACA 484
DB 452 GGATGTACAAACATGAAGACACTCTCTGTTGTCGCTTCTATTTATGACACCAACAGCT 511
QY 485 ACAGCAACAGCTTGATGATGTTCCCGAGAGATTCATTTGGCTGGGAAACGAGGCAAG 544
DB 512 TGGAGAACCGAGCTCGACCGATGCTTAGGAGATCTATCTGCTGGGAAACGAAGAGCAAG 571
QY 545 AGTTCTTAAGGTACACAGCAACAAAGCAGACAAAGCAGACGAGAGCTTACCATATAGCC 604
DB 572 AGTTCTTAAATATACGAAGAG----- 594
QY 605 CATACAGCCCGCAAGTCAGCCTTAGACAAGAGCGGTGAATTTAGCCCTTCGAGGAGCAGC 664
DB 595 -----CAAGAGGTATC 607
QY 665 ACAGCCGAGAGAACGAGCAGGACAGAAAGAAACGAAGGTGAAACATCTTCAGCG 724
DB 608 AAAGCCAGAAAGGAAAGCATCAGCAAGAAAGAAACGAAGGAGCAGCATATTGAGTG 667
QY 725 GCTTCAGCGCGGAGTTCCTGGAACAGCCTTCAGGTTGACGACACAGATAGTCGAAA 784
DB 668 GCTTCACCTCGGAATTTCTTGGAAACATGCAATTCAGCGT---GGCAAGCAGATAGCGAAA 724
QY 785 ACCTAAGAGCGAGACCGGAGGTGAAGAAAGAGGAGCCATTGTGACAGTGCAGGGAGGCC 844
DB 725 ACCTAAGAGGAGAACGAGAGGAGGAGAGCAAGGGAGCCATTGTGACAGTGAAGAGGATC 784
QY 845 TCAGAACTTTGAGCCCA-----GATGAAAGAGAGCGTCCGACGAGAGAGAGAAAT 895
DB 785 TGAGCGTGATAAAACCAACCGCAGCAGCAACAAAGACCCAGGAGAGAGAGAAAG 844
QY 896 ACGATGAAGTGAATG-----TGAAATACGATGAAGAGGATAGAGCGGTG 940
DB 845 AAGAGAGGATGAGAAAGCCACAGTGCAAGGTGAAGACAAACACTCCCAACGCCCCGAG 904
QY 941 GCAGGGAAAGCAGAGCGGAGGAGTGTATTTGAAGAGACGATCTGCACCGCAAGTGCTA 1000
DB 905 GAAGCCAAAGCAAAAGCAGAGAAATGGCATTTGACGAGACCATATGACCATGAGACTTC 964
QY 1001 AAAAGAAACATTTGGTAGAAAACAGATCCCTGACATCTTACAAACCTCAAGCTGGTTCACTCA 1060
DB 965 GCCAACAACTTTGGCCAGACTTCATCACCTGACATCTTACAAACCTCAAGCGGTAGCGTCA 1024
QY 1061 AACTGCCCAACGATCTCAACCTTCTAATCTTAGTGGTGGCTTGGACCTAGTCTGATATG 1120
DB 1025 CAACGCCCAACGCTTGAATCTTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTG 1084
QY 1121 GAAATCTCTACAGGAATGCAATTTGTCGCTCACTTACAAACCAACGACACGACGATCA 1180
DB 1085 GATCTCTCCGCAAGATGCAATTTGTCGCCACACTTACAACTGAAACGCAAGCATATA 1144
QY 1181 TATATCGATTGAGGGAGCGGCTCAGCTGCAAGTCTGTGGACAGCAACGCAACAGAGTGT 1240
DB 1145 TATACGCAATTGAATGCAAGCGGCATTGATACAAAGTGTGAAATGCAACGGTGAGAGAGTGT 1204
QY 1241 ACGAGAGAGCTTCAAGAGGGTCACTGCTGTGTGTGTCACAGAACTTTGCGCGTGCCTG 1300
DB 1205 TTGATGAGAGAGTGCAGAGGGAGCGGTGCTGATGCTGTCACAAAACCTTTGTGTGGTGG 1264
QY 1301 GAAAGTCCAGAGCAGAGAACTTCGAATAGTGTGCAATTTCAAGACAGACTCAAGGCCAGCA 1360
DB 1265 CAAGATCAGAGAGTGACAACTTCGAGTATGTGTCATTCAAGACCAATGATACACCCATGA 1324

```

QY 1361 TAGCCAACTCGCGGTGAAAACTCCGTCATAGATAAACCCTGCGGAGGAGTGGTTGCAA 1420
Db 1325 TCGCACTCTTCGAGGGGCAAACTCATTTGTGAACGATTTACACAGGAGTGAATTCAGC 1384
QY 1421 ATTATATGCGCTCCAAAGGAGAGCGCAAGCAGCTTAAGAAACAACCCCTTCAAGT 1480
Db 1385 ACACTTTCAACCTAAAAGGCGAGCGGCCACGCGCAGGAGATTAAGAAACAACCCCTTCAAGT 1444
QY 1481 TCTTGGTTCCACCGTCTCAGCAGTCTCCGAGGCGTGTGGCTTA 1523
Db 1445 TCTGGTTCCACCTCAGGAGTCTCAGAAAGAGCTGTGGCTTA 1487

RESULT 7
US-09-762-381-1
; Sequence 1, Application US/09762381
; Patent No. 6576820
; GENERAL INFORMATION:
; APPLICANT: Takaiwa, Fumio
; APPLICANT: Utsumi, Shigeru
; APPLICANT: Katsube-Tanaka, Tomoyuki
; TITLE OF INVENTION: TRANSGENIC PLANT EXPRESSING SOYBEAN GLYCININ
; FILE REFERENCE: 201487/1050
; CURRENT APPLICATION NUMBER: US/09/762,381
; PRIOR FILING DATE: 1999-03-04
; CURRENT APPLICATION NUMBER: PCT/JP99/01057
; PRIOR FILING DATE: 1999-03-04
; PRIOR APPLICATION NUMBER: JP 10-223-897
; PRIOR FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1743
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(1536)
US-09-762-381-1

Query Match      33.7%; Score 513; DB 4; Length 1743;
Best Local Similarity 62.7%; Pred. No. 2.5e-143;
Matches 968; Conservative 0; Mismatches 425; Indels 150; Gaps 5;

QY 5 AGCAACGGAGGAGAACGCGCTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTTGACAATC 64
Db 122 AGCAGCCTCAGCAAAACGAGTGCAGATCCAAAACCTCAATGCGCTCAAAACCGGATTAAC 181
QY 65 GCATTGAATCAGAGGCGGTTACATTGAGACTTTGGAACCCCAACCAACAGGAGTTCGAAT 124
Db 182 GTATAGAGTCAGAAGGAGGCGCTCATTTGAGACATGGAACCCCTAACAAAGCCATTCCAGT 241
QY 125 GCGCGCGGCTGCGCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCGCTTCT 184
Db 242 GTGCGCGGTGTGCGCTCTCTCGCTGACCCCTCAACCGCAACGCGCTTCGTAGACCTTCCT 301
QY 185 ACTCCAACTGCTCCAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTTGATAT 244
Db 302 ACACCAACGCTGCCAGGAATCTACATCCAAAGGTAAGGGAATTTTGGCATGATAT 361
QY 245 TCCTGTGTTGCTCTAGACACTATGAAGAGCGCTCACACACAAGTCCGTGATCTCAGTCCC 304
Db 362 ACCGGGTGCTCTAGCACATTTGAGAGCGCTCA----- 395
QY 305 AAAGACCAACAGAGCTCTCAAGGAGAGAGACCAAGACCAAGCAACGAGATAGTCAAC 364
Db 396 -----ACAACTTCAAAAGAGGAGCAAAAGCAGCAGACCAAGACCGTCACC 442
QY 365 AGAAGGTGCAACGCTTTCGATGAGGAGTATCTCATTCAGTTCCACCGGTTGCTTCT 424
Db 443 AGAAGATCTATACTTCAGAGGGGTGATTTGATCGCAGTGCCTACTGTTGTTGATGGT 502

```

```

QY 425 GGCTCTACAAACGACCAACGACACACACACACACACACACACACACACACACACACAC 484
Db 503 CGATGTACAAACATGAAGACACCTCTGTTGTTGCGCTTCTATTTATGACACCAACAGCT 562
QY 485 AGCAACACAGCTTGTATGATTTCCAGGAGATTTCAATTTGGCTGGGAACACGAGCAAG 544
Db 563 TGGAGAACACAGCTCGACCAAGATGCGCTAGGAGATTTATCTTCTGGGAACCAAGAGCAAG 622
QY 545 AGTTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACAAAGAGCTTTACCATATAGCC 604
Db 623 AGTTCTTAAATATCAGCAAGAG----- 645
QY 605 CATACAGCCCGCAAAAGTTCAGCCTTAGACAAAGAGCGTGAATTTAGCCTTCGAGGACAGC 664
Db 646 -----CAAGAGGTCATC 658
QY 665 ACAGCCGACAGAGAACGAGCAGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
Db 659 AAAGCCAGAAAGGAAAGCATCAGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY 725 GCTTACAGCGCGGAGTTCTTGGAAACAAAGCTTCCAGGTTGACGACACAGACAGATAGTCAAA 784
Db 719 GCTTACACCTGGAATTTCTTGGAAACATGCAATTCAGCGT---GGACACAGCAGATAGCGAAA 775
QY 785 ACCTAAGAGCGCAGACCCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAAGGAGGCC 844
Db 776 ACCTAAGAGGAGAGAACGAAAGGGAAGACAAAGGAGGCCATTGTGACAGTGAAGAGAGTGC 835
QY 845 TCAGATCTTGTAGCCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAGAGAT 895
Db 836 TGAGCGTGATAAACCAACCCACGAGCAGACGACAAAGAGACCCGAGGAGAGAGAGAGAG 895
QY 896 ACGATGAAGATCAATA-----TGAATACGATGAAGAGGATAGAAGCGCTG 940
Db 896 AAGAAGAGATGAGAAAGCCACAGTGCAGGGTTAAGACAAACACTGCCCAACGCCCGAG 955
QY 941 GCAGGGAGAGCAGAGGCGAGGGGAATGATTTGAAGAGAGAGATCTGCACCCGCAAGTGTCTA 1000
Db 956 GAAGCCAAAGCAAAAGCAGAGAAATGGCATTTGACGAGACCATATGACCATGAGACTTC 1015
QY 1001 AAAGAAACATTGCTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGTTCACTCA 1060
Db 1016 GCCACAACATTTGGCAGAGCTTCACTACCTGACATCTACAAACCTCAAGCCGTTAGCGTCA 1075
QY 1061 AAAGTCCCAACCATCTCAACCTTTAAATACCTTAGGTGGCTTGGACCTAGTGTGATATG 1120
Db 1076 CAACGCCACACAGCTTGACTTCCAGCCCTCTCGTGGCTCAGACTCAGTGTGAGTTG 1135
QY 1121 GAAATCTCTACAGGAATGATTTGTTGTCGCTCACTAACAACCAACGACACAGCATCA 1180
Db 1136 GATCTCTCCGCAAGAAATGCAATGTTCTGTCGCCACACTACAACTGAAACGCAACAGCATAA 1195
QY 1181 TATATCGATTGAGGGAGCGGCTCAGTGCAGAGTCTGTCGACAGCAACGCAACAGAGTCT 1240
Db 1196 TATACGATTTGAATGAGCGGCAATGATACAAAGTGGTGAATTTGCAACGCTGAGAGAGTGT 1255
QY 1241 ACACGAGAGCTTCAAGAGGGTTCAGTGTCTTGTGTGTCACAGAACTTTCGCGCTCGCTG 1300
Db 1256 TTGATGGAGAGCTGCAAGGAGCGGCTGCTGATCGTGCACAAAACCTTTGTTGCTGCTG 1315
QY 1301 GAAAGTCCCAAGCAGAGAACTTCGAAATACGTGGCAATTTAAGAGACAGACTCAAGGCCAGCA 1360
Db 1316 CAAGATCAACAGAGTGAACAACTTCGAGTATGTCTATTCAAGACCAATGATACACCCATGA 1375
QY 1361 TAGCCAACTCCCGGTGAAAACCTCCGTCATAGATAAACCCTGCGGAGGAGTGGTTGCAA 1420
Db 1376 TCGGCACACTTTGAGGGGGCAAACTCATTTGTTGAACGCAATTAACAGAGGAAGTGAATTCAGC 1435
QY 1421 ATTTCATATGCGCTTCAAAAGGAGGAGGCAAGGAGCTTAAGAAACAACCCCTTCAAGT 1480
Db 1436 ACACCTTCAACCTAAAAGCCAGAGCCAGGCGCAGATTAAGAAACAACCAACCCCTTCAAGT 1495
QY 1481 TCTTGTTCACCGCTCTAGCAGTCTCCGAGGGCTGTGGCTTA 1523

```

Db 1496 TCTGGTTCACCTCAGAGTCTCAGAGAGAGCTGGCTTA 1538

RESULT 8

US-09-108-010B-12
 ; Sequence 12, Application US/09108010B
 ; Patent No. 6362399
 ; GENERAL INFORMATION:
 ; APPLICANT: ANTHONY JOHN KINNEY
 ; GARY MICHAEL FADER
 ; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
 ; CLASSES OF SOYBEAN SEED
 ; PROTEIN GENES
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
 ; STREET: 1007 MARKET STREET
 ; CITY: WILMINGTON
 ; STATE: DELAWARE
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 19898

COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.50 INCH
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: MICROSOFT WINDOWS 95
 SOFTWARE: MICROSOFT WORD 7.0A
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/108.010B
 FILING DATE: 30-Jun-1998
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/019,940
 FILING DATE: JUNE 14, 1996
 ATTORNEY/AGENT INFORMATION:
 NAME: LYNNE M. CHRISTENBURY
 REGISTRATION NUMBER: 30,971
 REFERENCE/DOCKET NUMBER: BB-1071-A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 302-992-5481
 TELEFAX: 302-773-0164
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1458 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-108-010B-12

Query Match 33.1%; Score 504; DB 3; Length 1458;
 Best Local Similarity 62.0%; Pred. No. 1.2e-140;
 Matches 943; Conservative 0; Mismatches 450; Indels 129; Gaps 4;

Qy 5 AGCAACGGAGGAGACCGCTGCCAGTTCAGGCGCTCAATGCCAGAGACCTGACATC 64
 Db 62 AGCAGGACAGCAAAATGAGTGCAGATCCAAAAGCTGAAATGCCCTCAAAACCGGATAAC 121
 Qy 65 GCATTGAATCAGAGGCGGTTACATTCAGACTTGGAAACCCCAACCAACGAGGTTGCAAT 124
 Db 122 GTATAGAGTGGAGGTGGGTTCAATTGAGCATGGAACCCCTTAACCAAGCCATTCCAGT 181
 Qy 125 GCGCCGCGTGGCCCTCTCTCGCTTAGTCTTCGCGCCGCAACGCCCTTCGTAGGCTTTCT 184
 Db 182 GTGCGGCTGTGGCCCTCTCTCGCTGACCCCTTAACCGCAATGCCCTTCGTAGACCTTCT 241
 Qy 185 ACTCCAAATGCTCCCGAGGATCTTCATCAGCAGGAGGAGTACTTTGGGTGATAT 244
 Db 242 ACACCAACGCTCCCGAGGAAATCTACATACAAAGGTAATGTTATTTTGGCATGATAT 301
 Qy 245 TCCTGTTGCTCTAGACACTATGAAGAGCTCACACACAAAGGTTCGTGATCTCAGTCCC 304

302 TCCCGGGTGTCTAGCAGCTTATCAAGAGCCGCA----- 335
 Qy 305 AAAGACCAACAGAGCGTCTCAGAGGAGAGAGCAAGAGCCAAAGCAGCAGAGATAGTCACC 364
 Db 336 -----AGAAATCTCAGCAACGAGGAGCAAGAGGCGCCCAAGAGCGTCACC 382
 Qy 365 AGAAGGTGACCGCTTTTCGATGAGGGTGATCTCATTCGAGTTCCACCGGTTGCTTTCT 424
 Db 383 AAAAGGTACATCGCTTCAGAGAGGGTGATTTGATCGCAGTGCCCTACTGTTGGTTCATGGT 442
 Qy 425 GGCTCTACACAGCAGCAGCAGCAGCTGATGTTGTTGCTGTTCTTCTTACTGACCAACAACA 484
 Db 443 GGATGTACAAATGAAGACACTCTCTGTTGTTGCTGTTCTATTATTGACACCAACAGCT 502
 Qy 485 ACAGAAACGAGCTTGATCAGTCTCCAGAGAGATCAATTTGGCTGGGAAACAGGAGCAAG 544
 Db 503 TGGAGAAACAGCTCGACAGATGCTAGGAGATTTCTATCTTGTGGGAAACAGAGCAAG 562
 Qy 545 AGTTCTTAAGGTACACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC 604
 Db 563 AGTTCTTAATAATATCAGACGAG----- 585
 Qy 605 CATACAGCCCGCAAGTTCAGCTTAGACAAGAGCGTGAAATTTAGCCCTCGAGGACAGC 664
 Db 586 -----CAGCAAGGAGGTT 598
 Qy 665 ACAGCCGAGAGAAACGAGCAGGAGCAAGAAAGAAACGAGGTGGAAACATCTTCAGCG 724
 Db 599 CCAAGAGCCAGAAAGGAAAGCAACAAAGAAAGAAAGAAAGCAACATATTTAGTG 658
 Qy 725 GCTTCACGCGGAGTTCTGGAAACAGCTTCCAGGTTGACGACAGACAGATAGTGAAA 784
 Db 659 GCTTCGCCCTGTAATTTCTGAAAGAGCGTTCT---GGGTGAACATGCAGATAGTGAAA 715
 Qy 785 ACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGGCATTTGTGACAGTGGAGGAGGCC 844
 Db 716 ACCTAAGGTGAGAAAGAGAGAGGAGGATAGTGGAGCCATTGTGACAGTGAAGAGGAGTC 775
 Qy 845 TCAGAAATCTTGGAGCCAG---ATAGAAAGAGAGCTGCCGAGCAGAAAGAGAGAAATACGATG 901
 Db 776 TAAGAGTCACAGCTCCAGCCATGAGGAAAGCCACAGCAAGAAAGAGATGATGATGAGG 835
 Qy 902 AGATGAATATGAATACGATGAGAGAGGAGGAGGAGTGCAGGGAAGCAGAGGAGGAGG 961
 Db 836 AAGAGAGCCACAGTGGGTGAGAGACAGAAAGGTTGCCAACGCCCAAGCAAGAGGAGCA 895
 Qy 962 GGAATGGTATTGAAGAGAGAGCATCTGCACCGCAAGTGTCTAAAAAGAACATTTGGTAGAAACA 1021
 Db 896 GAAATGGCATTGATGAGACCATTTGCAATGAGACTTCGCCAAACATTTGGTCAGAAATT 955
 Qy 1022 GATCCCTGACATCTACAAACCTCAAGCTGGTTTCACTCAAAACTGCCAACGATCTCAACC 1081
 Db 956 CATCAGCTGACATCTACAAACCTCAAGCTGGTAGCATCAACCCGCCACAGCCTTGACT 1015
 Qy 1082 TTCATATCTAGTGGCTTGGACCTAGTGTCAATATGGAATCTCTACAGGAATGCAAT 1141
 Db 1016 TCCAGGCCCTCTGGCTTCTCAAACTCAGTGGCCAGTATGGATCACTCCGCAAGAAATGCTA 1075
 Qy 1142 TGTTCGTGCGTCTCACTACAAACCAACGAGCAGCAGCATCATATATTCGATTCAGGGGAGCGG 1201
 Db 1076 TGTTCGTGCCACACTACACCTGAAACGCGAAGACGATATATACGCAATGAAATGGGCGG 1135
 Qy 1202 CTCAGCTGAGTGTGGAGCAGCAACGAGGCAACAGAGTGTACGAGAGGAGCTTCAGAGAG 1261
 Db 1136 CATTTGGTCAAGTGGTGAATTTGCAATGCTGAGAGAGTGTGTTGATGGAGAGCTGCAAGAG 1195
 Qy 1262 GTCAGCTGCTTGGTGGCCACAGAACTTCGCCCTCGCTGGAAAGTCCACAGAGGAGAACT 1321
 Db 1196 GAGGGTGTGATGCTGTTCCAAACCTTTGCGGTGGCTGCAAAATCCAGAGGAGTAACT 1255
 Qy 1322 TCGAAATACGTGGGCTTCAAGAGACAGACTCAAGGCGCCAGCAGTACAGCAACCTCGCCGGTGA 1381
 Db 1256 TTGAGTATGTGCTATTCAAGACCAATGATAGACCTTCGATCGGAACCTTCGAGGGGCA 1315

Qy	1382	ACTCCGTCATAGATAA	CTCGCGAGGAGGTGGTTGCAAAATTATATGGGCTTCAAAGGG	1441
Db	1316	ACTCATTGTTGAACGCAT	TGCGAGGAAAGTGATTTCAGCACACTTTTAAACCTTAAGAGCC	1375
Qy	1442	AGCAGGCAAGGACGCTT	AAGAAACAACCCCTTCAAGTTCTTCGTGTCCACCGTCTCAGC	1501
Db	1376	AGCAGGCGAGGAGCTG	AGAGNACAAACCCCTTTCAGCTTCTGTGTCCACCTCAGGAGT	1435
Qy	1502	AGTCTCCGAGGCTGTG	CGCTTA	1523
Db	1436	CTCAGAGGAGAGCTGT	GCCTTA	1457
RESULT 9				
US-09-758-652-12				
; Sequence 12, Application US/09758652				
; Patent No. 6703544				
; GENERAL INFORMATION:				
; APPLICANT: ANTHONY JOHN KINNEY				
; APPLICANT: GARY MICHAEL FADER				
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC				
; CLASSES OF SOYBEAN SEED				
; PROTEIN GENES				
; NUMBER OF SEQUENCES: 21				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY				
; STREET: 1007 MARKET STREET				
; CITY: WILMINGTON				
; STATE: DELAWARE				
; COUNTRY: UNITED STATES OF AMERICA				
; ZIP: 19898				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: DISKETTE, 3.50 INCH				
; COMPUTER: IBM PC COMPATIBLE				
; OPERATING SYSTEM: MICROSOFT WINDOWS 95				
; SOFTWARE: MICROSOFT WORD 7.0A				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/09/758,652				
; FILING DATE: 11-Jan-2001				
; CLASSIFICATION: <Unknown>				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: 60/019,940				
; FILING DATE: JUNE 14, 1996				
; ATTORNEY/AGENT INFORMATION:				
; NAME: LYNNE M. CHRISTENBURY				
; REGISTRATION NUMBER: 30,971				
; REFERENCE/DOCKET NUMBER: BB-1071-A				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 302-992-5481				
; TELEFAX: 302-773-0164				
; INFORMATION FOR SEQ ID NO: 12:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1458 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: single				
; TOPOLOGY: linear				
; MOLECULE TYPE: cDNA				
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:				
US-09-758-652-12				
Query Match 33.1%; Score 504; DB 4; Length 1458;				
Best Local Similarity 62.0%; Pred. No. 1.2e-140;				
Matches 943; Conservative 0; Mismatches 450; Indels 129; Gaps 64				
Qy	5	AGCAACCGGAGGAGAA	CGGTGCGAGTTCCAGCGCCTCAATGGCAGAGACCTGACAAATC	64
Db	62	AGCAGGCACAGCAAAAT	GAGTGCCAGATCCAAAGCTGAATGCCCTCAAACCGGATTAACC	121
Qy	65	GCATTGTAATCAGAGG	CGGTTTACATTGAGACTTTGGAAACCCCAACCAACGAGGATTCGAAT	124
Db	122	GTAATAGAGTCGGAAG	TGGGTTTCATTGAGACATGGAACCCCTTAAACAACGACATTCAGT	181

125	Qy	GC	CGCGCGT	CGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCGCCTTCGTAGCGCTTTCT	184
182	Db	GT	CGCGGTGTGCGCCCTCTCTCGCTGCACCCCTTAACCGCAATGCCCTTCGTAGACCTTCCT	241	
185	Qy	ACT	CCAAATGCTCCCGCAGGAGATCTTCATCCAGCAAGAAAGGGGATACTTTGGGTTCGATAT	244	
242	Db	AC	ACCAACGGTCCCGAGGAACTACATACACAAGGTAATGGTATTTTGGCATGATAT	301	
245	Qy	TCC	CTGGTTGCTTAGACACTATGAAGAGCCTCACACAAGGTCGTCTGATCTCAGTCCC	304	
302	Db	TCC	CGGTTGCTTAGCACTTCTAGCACTTCTCAAGACGCGCA-----	335	
305	Qy	AA	AGACCAACAGACGCTCCAGGAGAGACCAAGCCAAACAGCAACGAGATAGTCAACC	364	
336	Db	-----	AGAACTCTAGCAACGAGGACGAAGCCAGAGGCCCAAGACCGTCAACC	382	
365	Qy	AGA	AGGTGACACCGTTTTCGATGAGGGTGATCTCATTTGCAGTTCACACCGGTGTGCTTTCT	424	
383	Db	AAA	AGGTACATCGCTTCAGAGAGGGTGATTTGATCGCAGTGCTACTGTGTGTGCATGGT	442	
425	Qy	GG	CTCTAACACGACACGACATGATGTTGTGTGCTGTTCTTCTTACTTGACACCAACAACA	484	
443	Db	GG	ATGTACAAACAATGAAGACACTCTCTGTGTGTGCCGTTCCTATATTATTTGACCAACAGCT	502	
485	Qy	AC	GACAAACAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAAACACGAGCAAG	544	
503	Db	TG	AGAACACAGCTCGACACAGATGCCCTAGAGATTCTATCTTCTGGGAAACCAAGACCAAG	562	
545	Qy	AG	TTCTTAAGGTACGACGACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATATAGCC	604	
563	Db	AG	TTTCTTAAATATCAGCAGCAG-----	585	
605	Qy	CAT	ACAGCCCGCAAGTCAAGCCTTAGACAAGAGAGCGTGAATTTTAGCCCTTCGAGGACAGC	664	
586	Db	-----	-----CAGCAAGGAGTT	598	
665	Qy	AC	AGCCGACAGAAACGAGCAGGACAAAGAGAAAGAAACGAAAGTGGAAAACATCTTCAGCG	724	
599	Db	CC	CAAGCCAGAAAGAAAGCAACAAAGAAAGAAACGAAAGGAAGCAACATATTGAGTG	658	
725	Qy	GC	TTCAAGCGGAGTTCTTGGAACAGCCCTCCAGGTTGACGACACAGACAGATAGTGCAAA	784	
659	Db	GC	TTGCGCCCTGAAATTTCTTGAAGAAAGCGCTC---GGCGTGAAACATGCGAGATAGTGAAGAA	715	
785	Qy	AC	CTAAGAGCGAGACCGCAGAGTGAAGAAAGAGGGAGCCATTGTGACAGTGAAGGGGAGGCC	844	
716	Db	AC	CTACAAGGTGAGACGAGAGGAGGATAGTGGAGCCATTGTGACAGTGAAGAGGAGTCT	775	
845	Qy	TC	AGAAATCTTGAGCCGAG---ATAGAAAGAGAGCTCCCGACGAAGAAAGGAATACGATG	901	
776	Db	TA	AGATCAAGCTCCAGCCATGAGGAAGCCACAGCAAGAAAGAAATGATGATGATGAGG	835	
902	Qy	AGA	TCAATATGAATACGATGAAGAGGATAGAAGGCGTGGCAGGGGAAGCAGAGGGCAGGG	961	
836	Db	AA	GACAGCCACAGTCCGTGGAGACAGACAAGAGTTGCCAACGCCAAGCAAAAGAGGCA	895	
962	Qy	GG	AATGGTATGAAGAGACGATCTGCACCGCAAGTGCTAAAAAGAAACATTTGGTAGAAAACA	1021	
896	Db	GAA	ATGGCATTTGATGAGACCAATTTGCACAATGAGACTTCGCCCAAAACATTTGGTCAGAAAT	955	
1022	Qy	GAT	CCCTGACATCTTACACCTCAGCTGGTTCACTCAAAACTGCCACGATCTCAACC	1081	
956	Db	CAT	CACTGACATCTTCAACACCTCAAGCTGGTAGCATCAACCCGCCACCGCCTTGACT	1015	
1082	Qy	TT	CTAATACTTAGTGGCTTGGACCTTAGTGCTGCAATATGGAATCTCTCAGGAAATGCAAT	1141	
1016	Db	TCC	AGCCCTCTGGCTTCTCAAACTCAGTGCCCGATGATCGATCACTCCGCAAGAAATGCTA	1075	
1142	Qy	TG	TTGTGTGCTACTTACAACAACCAACGACACAGCATCATATATCGATTTAGGGGAGCGGG	1201	
1076	Db	TGT	TCGTGCCACTACACCTGAAACGCCAAACAGCAATAATATACGATATTGAATGGCGCGG	1135	
1202	Qy	CT	CAGTGCAGGTCGTGGGACGCAACCGGCAACAGAGTGTTACGACAGGAGGAGCTCAAGAGG	1261	

Db 1136 CATTGGTACAGTGGTGAATTCGATGAGAGAGTGGTGGAGAGCTGCAAGAGG 1195
Qy 1262 GTCACTGCTGTGGTGGCCACAGAACTTCGCGCTGCTGGAAAGTCCCGAGAGCGAGAACT 1321
Db 1196 GAGGGTGTCTGATGTTCCCAAAACCTTCGGGTGGCTGCAGAAATCCCGAGCGATTAAT 1255
Qy 1322 TCGATACGTGGCATTCAAGACAGACTCAAGGCCCGAGCAGTAGCCAACTCGCGGTGAAA 1381
Db 1256 TTGAGTATGTCTATTCAAGACCAATGATAGACCTCGATCGGAAACCTTCGAGGGGCAA 1315
Qy 1382 ACTCCGTCATAGATAACTGCGGAGGAGGTGGTTGCAAAATTCATATGGCTCCAAAGGG 1441
Db 1316 ACTCATTTGTAACGCATTCCAGAGGAGTGAATTCAGCACATTTTAACTTAAAGAGCC 1375
Qy 1442 AGCAGGCAAGCGAGCTTAAGAACAAACCCCTTCAAGTTCTTGGTTCCACCGTCTCAGC 1501
Db 1376 AGCAGGCCAGCGAGTGAAGAACAAACACCTTTTCAGCTTCTTGGTTCCACCTCAGGAGT 1435
Qy 1502 AGTCTCCGAGGGTGTGGCTTA 1523
Db 1436 CTCAGAGGAGAGTGTGGCTTA 1457

RESULT 10

US-10-684-651-12
; Sequence 12, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; APPLICANT: GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEWMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNN M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1458 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-684-651-12

Query Match 33.1%; Score 504; DB 4; Length 1458;
Best Local Similarity 62.0%; Pred. No. 1.2e-140;
Matches 943; Conservative 0; Mismatches 450; Indels 129; Gaps 4;
Qy 5 AGCAACCGGAGGAGAACGGGTGCGAGTTCCAGCGCTCAATGCGCAGAGACCTTGACAATC 64
Db 62 AGCAGGCAACGCAAAATGAGTGCAGATCCAAAAGCTGAATGCCCTCAACCCGATNACC 121
Qy 65 GCATTGAATCAGAGGCGGTTCATTGAGACTTTGGAAACCCCAACAAACAGAGAGTTCGAAT 124
Db 122 GTATAGAGTCGGAAGGTGGGTTTCATTGAGACATGGAACCTTAACAACAGCCATTCAGT 181
Qy 125 GCGCCGGGTGCGCCCTCTCGCTTAGTCTCGCGCGCAGCGCCCTTCGTAGGCTTCT 184
Db 182 GTGCCGGTGTGCCCTCTCTCGCTGACCCCTTAAACCGCAATGCCCTTCGTAGAGCTTCT 241
Qy 185 ACTCCAAATGTCCTCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTGATAT 244
Db 242 ACACCAACGGTCCCGAGGAAATCTACATACACAGGTAAATGATATTTTGGCATGATAT 301
Qy 245 TCCTCGTTGTCTTAGACACTATGAAGAGCTTCACACAAAGGTCTCGATCTCAGTCCC 304
Db 302 TCCTCGTTGTCTTAGACACTTATCAAGAGCGCA----- 335
Qy 305 AAAGACCAACAGAGCGTCTCCAAGGAGAGAACCAAGCCAAACAGCAACGAGATAGTCACC 364
Db 336 -----AGAACTCTCAGCAACGAGGACGAGAGCGCCCAAGACCGTCACC 382
Qy 365 AGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCCCGCTCCACCGGTGTTCTTCT 424
Db 383 AAAGGTACATCGCTTCAGAGAGGGTGATTTGATCCAGTGCCTACTGTGTTGCTATGGT 442
Qy 425 GGCTCTAACACGACCAACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAACAACA 484
Db 443 GGATGTACAAACAAATGAAGACACTCTCTGTTGTCGCTTCTATTAATTGACACCAACAGCT 502
Qy 485 ACAGAACAGCTTGATCAGTTCCTCCAGAGATTCATTTGGTGGGAACACGAGCAAG 544
Db 503 TGGAGAACAGCTCGACCAAGTCTCAGAGATTTCTATCTGCTGGGAACCAAGAGCAG 562
Qy 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATATAGCC 604
Db 563 AGTTCTTAATATCAGCAGCAG----- 585
Qy 605 CATAAGCCCGCAAGTTCAGCTTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGACAGC 664
Db 586 -----CAGCAAGGAGGTT 598
Qy 665 ACAGCCGAGAGAACGAGCAGGACGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
Db 599 CCAAGCCAG 658
Qy 725 GCTTCAGCGCGGAGTTCTCGGAACAGCTTCCAGGTTCAGCAGACAGACAGATAGTCAAA 784
Db 659 GCTTCGCCCTGAAATCTTGAAGAGAGCGGTTCC-----GGGTGAACATGCAATAGTAGAA 715
Qy 785 ACCTAAGAGCGGAGACCGAGAGGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGAGGCC 844
Db 716 ACCTAAGAGGTGAGAACGAAGAGAGGAGATAGTGAGGCCATTGTGACAGTGAAGAGGAGTC 775
Qy 845 TCAGAAATCTTGAGCCAG-----ATAGAAAGAGAGCTGCCAGCAGAGAGAGAGAGAGATGATG 901
Db 776 TAAGAGTCAAGCTCCAGCCATGAGGAGGCCACAGCAAGAGAGAGAGATGATGATGAGG 835
Qy 902 AAGATGAATATGAATACGATGAAGAGGAGTAGAAGGGGTGGCAGGGGAGAGAGAGAGG 961
Db 836 AAGAGCAGCCACAGTGGTGGAGACAGACAAAGTTGCCAAACGCCAAAGCAAGAGAGCA 895
Qy 962 GGAATGGTATTGAAGAGAGAGATCTGCAACCGCAAGTCTTAAAGAGACATTGGTAGAAACA 1021
Db 896 GAAATGGCATTGATGAGACCAATTGACAAATGAGACTCTGCCCAAAACATTGGTCAGAAAT 955
Qy 1022 GATCCCTCTGACATCTTACAAACCCCTCAAGCTGGTTTCACTCAAAAACCTGCCAACGATCTCAACC 1081

Db 956 CATCACTGACATCTAACCCTCAAGCTGGTAGCATCAACCGCCACCGCCTTGACT 1015
Qy 1082 TTCTAATACCTAGTGGCTGGACCTAGTCTCAATATGGAATCTACAGGAATGCAT 1141
Db 1016 TCCAGCCCTCTGGCTCTCAAACTCAGTGGCCCGATGATGATCACTCCGCAAGAATGCTA 1075
Qy 1142 TGTTCGTGCTCACTCAACACCAACGCGCACACAGCATCATATATCGAATGGAGGGACGGG 1201
Db 1076 TGTTCGTGCCACACTACACCTGAAACGCGAACAGCATATATACGCATTTGAATGGCGGG 1135
Qy 1202 CTCAGGTGCAAGTCTGGAGCAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGG 1261
Db 1136 CATTCGTACAAGTGGTGAATTCGAATGCTGAGAGAGTGTTCGATGAGAGCTGCAAGAGG 1195
Qy 1262 GTCACGTGCTTGGTGGCCACAGAACTTCGCCGTCTGCTGGAAGTCCAGAGCGAGAACT 1321
Db 1196 GAGGGGTGCTGATCGTTTCCAAAACTTTGCGGTGGCTGCAAAATCCAGAGGGATACT 1255
Qy 1322 TCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGGTGAAA 1381
Db 1256 TTGAGTATGTGTCATTCAAGACCAATGATAGACCTCTCGATCGGAACCTTGCAGGGGCAA 1315
Qy 1382 ACTCCGTATAGATAACCTCGCGAGGAGGTGTTGCAAAATTCATATGGCTCCAAAGGG 1441
Db 1316 ACTCATTTGTTGAACGCAATTCGCCAGAGGAGTGAATTCAGCACACTTTTAACTTAAAGAGCC 1375
Qy 1442 AGCAGCGAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGC 1501
Db 1376 AGCAGCGGAGGAGGTGAAGAACAAACCCCTTCAGCTTCTTGTTCACCTCAGGAGT 1435
Qy 1502 AGTCTCCGAGGGCTGTGGCTTA 1523
Db 1436 CTCAGAGGAGCTGTGGCTTA 1457

RESULT 11

US-09-108-010B-15
; Sequence 15, Application US/09108010B
; Patent No. 6362399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNN M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164

; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-108-010B-15

Query Match 13.0%; Score 198; DB 3; Length 1551;
Best Local Similarity 51.4%; Pred. No. 1.1e-48;
Matches 699; Conservative 0; Mismatches 60; Indels 61; Gaps 8;

Qy 14 AGGAGAACGCGTCCAGCTTCCAGCGCTCAATGCCAGAGACCTGACAAATCCGATTGAAT 73
Db 86 AGTTCAACGAGTGCCTCACTCAACACCTCAAGCGCTTGAACCCGACCCGCGTTGAT 145
Qy 74 CAGAGGGCGGTTTACATTGAGACTTGGAAACCCCAACACAGGAGTTTGAATGCGCGGG 133
Db 146 CGAAGGTGGTCTTATTGAACAATGAACTCTCAACACCCCTGAGCTGCAATGCGCGGTG 205
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCTTTCTACTCAATG 193
Db 206 TCACTGTTTCCAAACGCAACCCCTCAACCGCAACGGCTCCCACTTGCCATCTTTACTT 265
Qy 194 CTCGCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTGATTTCCCTGG 253
Db 266 ATCCCAATGATCATGTGTTCAAGGGAAGGAGCAATGGAATTCGATTTCCGGAT 325
Qy 254 GTCTTAGACACTATGAAGAGCTCAACAAAGGTGTCGATCTCAGTCCCAAGAACAC 313
Db 326 GTCCGAGACGTTTGAGAAGCCACA-----ACAAAC 355
Qy 314 CAAGAGCTCTCAAGAGAGAACCAAGCCACAGCAACAGATAGTACACAGAAGTGC 373
Db 356 AATCAAGCAAGAGAGGCTCAAGGTCAAGAGCACTACAGACAGTACCAAGAGATTC 415
Qy 374 ACCGTTTCGATGAGGCTGATCTCATTTGCAGTTCCCAACCGGTGTTGCTTTCTGGCTT 433
Db 416 GTCACTTCAATGAAGAGAGAGCTACTAGTATCTCTTGGTGTCTTACTTGACCTATA 475
Qy 434 AGCAGCAGACACTGATGTTGTTGTTTCTTCTACTGACACCAACAAACAGCAACC 493
Db 476 ACACCTGGCGATGAACCAAGTGTGTCATCAGTCTCTTGACACCTCCAACCTTCAACA 535
Qy 494 AGCTTGATCAGTTCCCGAGGATTCATTTTCGCTGGGAACACGCGAGCAAGAGTTCTTAA 553
Db 536 AGCTTGATCAAAACCCCGAGATTTTTCCTTGGTGGAAACCCAGATATAGAGCATCCCG 595
Qy 554 GGTACCCAGCAAAAGCAGCAAAAGCAGACGAAAGCTTACCATATAGCCCATACAGCC 613
Db 596 AGACCATGCAACA--CAGCAGCAGCAGAGAGTATGTTGGACGCAAGCGGGGCAACA 653
Qy 614 CGCAAGTCAGCTTACAAAGAG--AGCGTGAATTTAGCCCTCGAGGACAGCAGCCGC 672
Db 654 CGCAGCAGCAGGAGGAAGGTGGCAGTGTGCTCAGTGGCTTTCAGCAACATTTCTTAGC 713
Qy 673 AGAGAACGAGCAGGACAGAAAGAAACGAGGTGGAAACATCTTCAG----CGGCTT 728
Db 714 ACAATCTTCAACACCAACGAGGACACAGTGAAGAACTTCGGTCTCCAGATGACCAAG 773
Qy 729 CACGCGGAGTTCTCGGAACAAAGCCCTTCAGGTTGACGACAGACAGATAGTGCAAAACCT 788
Db 774 GAAGCAGATCGTGACAGTGGAGGGGCTCAGCGTTATCAGCCCAAGTGGCAAGACA 833
Qy 789 AAGAGCGAGACCGAGAGTGAAGAGAGGAGGCCATTTGACAGTGAAGGGGAGGCTCAG 848
Db 834 AGAAGACGAAGACGAAGACGAAGAGAAATATGAGCGGAGCCCTCTTATCTCTCAGC 893
Qy 849 AATCTTTGAGCCCGATAGAAAGAGAGCGTGC---GACGAAGAGAGGAATAAGTGAAGA 905
Db 894 ACACCAAGCCATGGAAAGCATGAAGATGACAGGACGAGGAGCAAGAGAGATCAACC 953

QY 906 TGAATATGAATAGATGAGAGATAGAAG-----GGTGGCAGGG 947
Db 954 TCGTCTGTATCAGCTCCACAGGACCAAGCAGGCGCCGGAACAAGAACACGCTGGAAG 1013
QY 948 AAGCAGAGGAGGAGGAGTATTAAGAGAGAGATCTGCACCGCAGAGTCTAAAGAA 1007
Db 1014 AGATGTCAGATAGAAATGGGTTGAGAAATATTTGACCATGAGCTTCACAGAA 1073
QY 1008 CATTTGGTAGAAACAGATCCCTGACATCTACAACTTCAAGCTGTTTCACTCAAACTGC 1067
Db 1074 CATTTGCTCGCCCTTCAGTCTGACTTCTACAACTTCAAGCTGTTTCACTCAAACTGC 1133
QY 1068 CAAGGATCTCAACTTCTAATAGTGGTCTGACCTAGCTAGCTGTAATGAAATCT 1127
Db 1134 CAACAGTCTCACTCTCCAGCCCTCCGCAATTCGAGCTCAGTGGCCCAATATGTTGCT 1193
QY 1128 CTACAGGAATGCAATGTTTCTGCTCTACTACAACTTCAAGCTGTTTCACTCAAACTGC 1187
Db 1194 CTACAGGAATGCAATGTTTCTGCTCTACTACAACTTCAAGCTGTTTCACTCAAACTGC 1252
QY 1188 ATTGAGGGAGCGGCTCACGTGCAAGTCTGAGCAGCAAGCAAGCAAGCAAGCAAGCA 1247
Db 1253 CTCGAGGGAAGG--AAGAGTAGAGTGTGAACTGCCAAGGGAATGCAAGTGTGAGCG 1310
QY 1248 GAGCTTCAAGAGGTCACGTGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1307
Db 1311 TGAGCTAAGGAGGAGCAATTTGCTAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1370
QY 1308 CCAGAGCGAGAACTTCAATACCTGCGATTCAGACAGAC 1347
Db 1371 GGGAGAACAGGATTTGAATATGTTAGTGTTCAGACACAC 1410

RESULT 12

US-09-758-652-15
; Sequence 15, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-758-652-15
Query Match
Best Local Similarity 51.4%; Pred. No. 1.1e-48;
Matches 699; Conservative 0; Mismatches 600; Indels 61; Gaps 8;
QY 14 AGGAGAACCGGTGCGAGCTTCCAGCGCTCAATGCGCAGAGACCTGACAAATCGCATTTGAAT 73
Db 86 AGTTCAACGAGTGCACCACTCAACACTCAACGCTTGGAAACCGGACACCGGTTGAGT 145
QY 74 CAGAGGGCGGTTTACATTTGAGACTTGGAAACCAACCAACAGGAGTTTCAATGCGCGGCG 133
Db 146 CCGAAGTGTGTTTATTTGAACATGAACTCTCAACACCTCTGAGCTGCAATGCGCGGTG 205
QY 134 TCGCCCTCTCTCGCTTAGTCTTCCGCGCAACGCCCTTCTGAGGCTTTCTTACTCCAATG 193
Db 206 TCACTGTTTCCAAACGACCTCAACCGCAACGGCTCCCACTTGCATCTTACTTACCTT 265
QY 194 CTCCCGAGGAGTCTTCACTCAGCAGGAGGAGGATCTTTGGTTGATATTCCTCTGGTT 253
Db 266 ATCCCAAAATGATCATTTGCTTCAAGGAGGAGGAGCAATTTGATTTGCAATTTCCGGAT 325
QY 254 GTCCTAGACACTATTAAGAGCCTCAACACAAAGGTCTGCTCTCAGTCTCCAAAGACAC 313
Db 326 GTCCTAGAGCTTTGAGAGCCACA-----ACAAAC 355
QY 314 CAGAGCTCTCAAGGAGAGCAAGCAAGCAACAGCAAGATATGTAACCAAGAGTGC 373
Db 356 AATCAAGCAGAGAGGCTCAAGGTCAAGCAGCAAGCACTAAGACAGCTCACCAGAGATTC 415
QY 374 ACGTTTTCATGAGGCTGATCTCAATGCAAGTTCCTCCACCGCTGTTGCTTCTGCTCTACA 433
Db 416 GTCCTTCAATGAGGAGAGCTACTAGTATTCCTCTTGGTTTCTTACTTGACCTATA 475
QY 434 AGACCCAGCAGACTGATGTTGTTGCTTCTTCTTACTGACACCAACAAACAGCAAC 493
Db 476 AACTGGCATGNAACAGTTGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 535
QY 494 AGCTTGATCAGTTTCCCAGGAGATTCATTTGGCTGGAAACACGAGGAGAGTTCTTAA 553
Db 536 AGCTTGATCAAAACCCAGAGTATTTTACCTGCTGGAAACCCAGATATAGAGCATCCCC 595
QY 554 GGTACAGCAACAAAGCAGACAAAGCAGAGAGAGCTTACCATATAGCCCATAGCC 613
Db 596 AGACCATGCAACAA--CAGCAGCAGCAGAGAGTCTATGTTGGACGCAAGCGGGCAACA 653
QY 614 CGCAAGTCAAGCTAGACAAAGAG--AGCGTGAATTTAGCCCTCGAGCAGACAGCGCG 672
Db 654 CGACAGCAGGAGGAGAGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 713
QY 673 AGAGAACGAGCAGGACAAGAAAGAAACAGAGGTTGAAACATCTTCTAG-----CGGCTT 728
Db 714 AATCTCTTCAACACCAACAGCAGCAGCTGAGAAACTTCGGTCTCCAGATCAGCAAG 773
QY 729 CAGCCGGAGTTCTTGGAAACAGCTTCCAGGTTTACGAGTACAGACAGATAGTGAACACCT 788
Db 774 GAAGCAGATCGTACAGTGGAGGAGGCTTACGCTTATCAGCCCAAGTGGCAAGACA 833
QY 789 AAGAGCGGAGACGAGAGTGAAGACAGGAGGAGCATTGTTGACAGTGAAGGAGGCTCAG 848
Db 834 AGAAGCAG 893
QY 849 AATCTTTGAGCCAGATAGAAAGAGAGCTGCC---GACGAAGAGAGAGAGAGAGAGAGAGAG 905
Db 894 ACGACCAAGCCATGGAAGAGCATGAAGATGACGAGGAGAGAGAGAGAGAGAGAGAGAGAG 953
QY 906 TGAATATGAATAGATGAAGAGATAGAAG-----CGGTGGCAGGG 947

Db 954 TCCTCCTGATCACCTCCACAGCGACCAAGCAGCGCCGCAACAACCAAGACCGCTGGAAG 1013
Qy 948 AACGAGGCGAGGGGNAATGTTAAGAGAGAGCATCTGCACCGCAAGTGTCTAAAGAA 1007
Db 1014 AGGATGTCAGACTAGAAATGGGGTTGAGGAAATATTTTGACCATGAAGCTTCACGAA 1073
Qy 1008 CATTGGTAGAAACAGATCCCTGACATCTACAACTCTCAAGCTGGTTCACTCAAAACTGC 1067
Db 1074 CATTGCTGCCCTTCAGTGTGACTTCTACAACTCCAAAGCTGGTGCATAGACCTT 1133
Qy 1068 CAACGATCTCAACCTTCTAATATTAGTGGCTTGGACCTAGTGTCTGAATATGAAATCT 1127
Db 1134 CAACAGTCTCACCTCCAGCCCTCCGCAATTCGGACTCAGTGCCTCAATATGTTGCTCT 1193
Qy 1128 CTACAGGAATGCAATGTTTGTCTCTACTACACCAACGACACAGCATCATATCG 1187
Db 1194 CTACAGGAATGGAATTTACTCTCAGATTGGAATTTGAACGCGAAACAGTGTGACG-ATGA 1252
Qy 1188 ATTGAGGGGACGGGCTCAGCTGCAAGTCTGTGACAGCAACGCGCAACAGAGTGTACGCGA 1247
Db 1253 CTCGAGGGAAGG--AAGATTAGAGTGGTGAATCTGCCAAGGAATGCAGTGTTCGACGG 1310
Qy 1248 GGAGCTTCAAGAGGTCACGTGTTGTGTGCCACAGAACTTCGCGTGCCTGGAAGTTC 1307
Db 1311 TGAGCTAAGAGGGGACAAATTGCTAGTGGTCCGCGAGAACCCCGCGGTGGCTGAGCAAG 1370
Qy 1308 CAGAGCGAGAACTTCGATACCTGCGCATTCAGACAGAC 1347
Db 1371 GGGAGAACAGGATTGGAATATGATAGTGTTCAGAGCACAC 1410

RESULT 13

US-10-684-651-15
; Sequence 15, Application US/10684651
; Patent No. 6828491
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; APPLICANT: GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-10-684-651-15

Query Match 13.0%; Score 198; DB 4; Length 1551;
Best Local Similarity 51.4%; Pred. No. 1.1e-48;
Matches 699; Conservative 0; Mismatches 600; Indels 61; Gaps 8;

Qy 14 AGGAGAACGCGTSCCGCTTCAGAGCGCTCAATGCGCAGAGACCTGACAAATGCATTTGAAT 73
Db 86 AGTTCAACGAGTGCCTCACTCAACAACTCAACGCGTTGGNACCCGACACCGCGTTGAGT 145
Qy 74 CAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTCGAATGCGCGCGG 133
Db 146 CGAAAGGTGCTTATTGAAACATGGAATCTTCAACACCTGAGCTGCAATGCGCGGTG 205
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCTGTTAGGCTTTCTACTCAATG 193
Db 206 TCACCTGTTTCCAAACGCAACCTCAACCGCAACGCGTCCCACTTGCCATCTTACTTACCTT 265
Qy 194 CTCCCAGGAGATCTTTCATCCAGCAAGGAAGGGGATACCTTTGGGTTGATATTTCCCTGTT 253
Db 266 ATCCCCAAATGATCATTTGCTTCAAGGGAAGGAGCAATTCGATTTGCAATTTCCGGAT 325
Qy 254 GTCTTAGACATATGAAAGACCTCAACAAAGTCTGTCATCTCAGTCCCAGAACCAAC 313
Db 326 GTCCCGAGACGTTTGAGAAGCCACA-----ACAAAC 355

Qy 314 CAAGACGCTCTCAAGGAGAGACCAAGCCACAGCAACGAGATAGTACACGAGAAGTGC 373
Db 356 AATCAAGCAGAAGAGGCTCAAGGTCAAGCAGCAACTACAAGACAGTACCCAGAAGATTC 415
Qy 374 ACCGTTTCGATGAGGCTGATCTCTCATTTGCAGTTTCCCAACGCGTGTGCTTTCTGGCTCTACA 433
Db 416 GTCACTTCAATGAAGGAGACGCTAGTGTATTCCTCTTGGTCTTCTTACTTGAACCTATA 475
Qy 434 ACGACAGACACTGATGTTGTTGCTGTTTCTCTTCTGACACCAACAAACAGCAACCC 493
Db 476 ACACTGGCGATGAACCAAGTTGTTGCCATCAGTCTCTTTGACACCTCCAACTTTCAACAATC 535
Qy 494 AGCTTTGATCAGTTCCCGCAGGAGATTCAAATTTGGCTGGGAAACACGCGAGCAAGAGTTCTTAA 553
Db 536 AGCTTGATCAAAACCCAGAGTATTTTACCTTGTCTGGGAAACCCAGATATAGAGCATCCCG 595
Qy 554 GGTACACGCAACAAAGCAGACAAAGCAGACGCAAGAGCTTTACCATATAGCCCATACAGCC 613
Db 596 AGACCATGCAACA--CAGCAGCAGCAGAAGAGTCTATGTTGGACGCAAGCAGGGGCAACA 653
Qy 614 CGCAAGTCAGCTTAGACAAGAAG-AGCGTGAATTTAGCCCTCGAGGACAGCAGACGCGC 672
Db 654 CGCAGCAGCAGGAGGAAGGTTGGCAGTGTGCTCAGTGGCTTTCAGCAACATTTCTTACG 713
Qy 673 AGAGAACGAGCAGGACAAAGAAAGAAACGAAAGGTGGAAACATCTTCTCAG----CGGCTT 728
Db 714 ACATCTCTTCAACACCAACGAGGACACAGCTGAGAAACTTCGGTCTCCAGATGACGAAG 773
Qy 729 CACGCGGAGTTTCTTGGAAACAGCCCTTCAGGTTTGACGACAGACAGATAGTGCAGAAACCT 788
Db 774 GAAGCAGATCGTGACAGTGGAGGGGCTCAGCGTTATCAGCCCAAGTGGCAAGAAACA 833
Qy 789 AAGAGCGGAGACCGAGAGTGAAGAAAGAGGAGCCATTTGTGACAGTGGGGGAGGCTCAG 848
Db 834 AGAAGACGGAAGCAAGAGCAGAGACGAAGAAATATGACGCGACGCGCTCTTATCTCTCCAG 893
Qy 849 AATCTTTGAGCCCGATAGAAAGAGAGAGCTGCC---GACGAAGAGAGGAATAAGTGAAGA 905
Db 894 ACGACCAAGCCATGGAAGAGCATGAAGATGAACGAGGACGAGGACGAGAAAGATCAACCC 953
Qy 906 TGAATATGAATACGATGAAGAGGATAGAAG-----GCGTGGCAGGGG 947

Db 954 TCGTCTGATCACCTCCACAGCCACCAAGCAGCGCCGGAACAACAAGAACACACGTGGAAG 1013
Qy 948 AAGCAGAGGAGGGGGAATGGTATTTGAAGAGACGATCTGCACCCGAAGTGTCTAAAAAGAA 1007
Db 1014 AGGATGTCAGACTAGAAATGGGTTGAGGAAATATTTGCACCATGAAGCTTCACGAGAA 1073
Qy 1008 CATTGGTAGAACAAGATCCCTGACATCTCAACACCTCAAGCTGTTCACCTCAAAACTGC 1067
Db 1074 CATTGCTCGCCCTTCACGTGCTGACTTCTCAACCCCAAAAGCTGTGCGCATTAGCACCCCT 1133
Qy 1068 CAACGATCTCAACCTTCTAATACCTAGTGTGCTTGGACCTAGTCTGAATATGAAATCT 1127
Db 1134 CAACAGTCTACCTCCAGCCCTCGGCAATTCGGAATTCGGAATTCGGAATTCGGAATTCG 1193
Qy 1128 CTACAGGAATGCAATGTTGCTGCTACTCAACAACCAAGCAACAGCAAGTGTACGAGCA 1247
Db 1194 CTACAGGAATGCAATGTTGCTGCTACTCAACAACCAAGCAACAGCAAGTGTACGAGCA 1252
Qy 1188 ATTGAGGGAGCGGCTCAGCTGCAAGTCTGGAAGTCTGGAAGTCTGGAAGTCTGGAAGT 1247
Db 1253 CTCGAGGAAAGG--AAGAGTTAGAGTGGTGAATTCGCAAGGGAATGCAAGTGTTCGACGG 1310
Qy 1248 GGAGCTTCAAGAGGCTCAGCTGCTGTTGGTGGCCACAGAACTTCGCCGTGCTGGAAGTC 1307
Db 1311 TGAGCTAAGGAGGGAACAATTGCTAGTGGTGGCCAGAACTTCGCCGTGCTGGAAGTC 1370
Qy 1308 CCAGAGGAGAACTTCGAATACGTTGGCAATTCGAAGCAGAC 1347
Db 1371 GGGAGAACAGGATGGGAATATGATGTTCAAGACACAC 1410

RESULT 14
US-09-108-010B-14
; Sequence 14, Application US/09108010B
; Patent No. 6162399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108.010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-108-010B-14
Query Match 10.1%; Score 153.2; DB 3; Length 1689;
Best Local Similarity 49.7%; Pred. No. 3.2e-35;
Matches 726; Conservative 0; Mismatches 578; Indels 156; Gaps 7;
Qy 14 AGGAGAACGGGTGCCAGTTCACAGCGCTCAATCGCAGAGACCTCGACAATTCGATTTGAAT 73
Db 83 AGCTCAACGAGTCCCACTCAACAACCTCAACGGTTGGAAACCCGACCAACGGGTTGAGT 142
Qy 74 CAGAGGGCGGTACATTTGAGACTTGGAAACCCCAACCAACAGGAGTTCGAATGCGCCGGC 133
Db 143 TCGAAGGTGGTTTGAATTTCAACAATGGAACCTCTCAACACCTCGAGCTGAAATGCGCCGGT 202
Qy 134 TCGCCCTCTCTCGCTTAGTCTCCGCGCAAGCCCTTCGTAGCGCTTTCTACTTCAATG 193
Db 203 TCACTGTTTCCAAACTCACCTCAACCGCAATGGCCTTCACTTGCATCTTTACTTCACTT 262
Qy 194 CTCCCGAGGAGATCTTTCATCCAGCAAGGAGGGGATACTTTGGGTTGATATTCCTCTGGTT 253
Db 263 ATCCCGGATGATCATCATCGCCCAAGGGAAGAGCACTGCAGTGCA---AGCCAGGAT 319
Qy 254 GTCTTAGACATATGAGAGCTTCAACAACAGGTGCTGATCTCAGTCTCCAAAGACCA 313
Db 320 GTCTGAGAGCTTTGAGGAGCCCAAGAACAA-----TCAACAAGAGAGGCTC 368
Qy 314 CAAGAGCTCTCAAGGAGAGAACCAAGCAACAGCAAGATGATCTACCAAGAGGTGC 373
Db 369 AAGGTGCG-----CAGAAGCAGCAGCTACAGGACAGTCCACAGAAGATTC 412
Qy 374 ACCGTTTCGATGAGGTGATCTCATTTGCAGTTCCCAACCGGTGTTGTTCTTGGCTCTACA 433
Db 413 GTCATCTCAATGAGAGAGAGCGTACTCGTGTCTCTCTGTTGTTCTTACTTGACCTATA 472
Qy 434 AGCACCACGACACTGATGTTGTTGTTGTTCTTCTTACTGACACCAACAACAGCAAC 493
Db 473 ACACCTGGCGATGAACCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 532
Qy 494 AGCTTGTAGTTCCTCCAGGAGATTCATTTGGCTGGACACGAGCAGAGTCTTAA 553
Db 533 AGCTTGTAGTTCCTCCAGGAGATTCATTTGGCTGGACACGAGCAGAGTCTTAA 592
Qy 554 GGTACCAAGCAACAAAGCAGACA----- 575
Db 593 AGACCATGCAACAACAACACAGCAAGAAAGTCAATGGTGGACGACAGCGGGGCAACACC 652
Qy 576 -----AAGCAGACGAAGAGCTTACATATAGCCCATACAGCCCGCAAAAGTCAAGCTAG 629
Db 653 AGCAGGAGGAAGAGGAAGAGTGGCAGCGTCTCAGTGGCTTTCAGCAACACTTCTTGG 712
Qy 630 ACAAGAGAGCGTGAATTTAGCCCTCGAGGACAGCAGCCGCGAGAGAAACATTCAGTCT 689
Db 713 CACAATCTTTCAACACCAACAGGACATAGCTGAGAAATTCAGTCTTCCAGACGACGAAA 772
Qy 690 AGAAGAGAAACGAAGGTGGAACATCT--TCAGCGCTTCACGCGGAGTTCCTGGAAC 748
Db 773 GGAAGCAGATCGTGACAGTGGAAAGAGGTCTCAGCGTTATCAGCCCTTCAAGTGGCAAG 832
Qy 749 AAGCCTTCCAGGTTGACGACAGACAGATAGTGC AAA----- 785
Db 833 AACAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 892
Qy 786 -----CCTAAGAGGAGACCGAGTGAAGAGAGGAGGACATTTGTACAGTAGGG 838
Db 893 CTCTCGCCGACCAAGCCATGGAAGCGTGAAACAAGACGAGGACGAGGACGAAGATGAAG 952
Qy 839 GAGGCTCAGAACTTTGAGCCAGATAGAAAGAGAGTGTCCGACGAGAGAGAGATACG 898
Db 953 ATAAACCTGCTCTAGTGGACCAAGCAAGCAAGCGTGAACAGAGCCAGGACGAGACG 1012

QY 899 ATGAAGATGAATATGATATCGATCAAGAGGATAGAAAGCGTGGCAGG----- 945
Db 1013 AGGACGAAGATGAAGATGAAGATCAACCTCGCAGAGCCGCGAATGGAGATCGAAAAAGA 1072
QY 946 -----GGAAGCAGGCGAGGGGAATG 967
Db 1073 CACAACCCAGAAGACCTTAGACAAGAAGAACACCGTGAAGAGGATGCGAGACAAGAAACG 1132
QY 968 GTATTGAAGAGAGATCTGCACCGCAAGTGTCTAAAGAGACATTTGTTAGAAACAGATCCC 1027
Db 1133 GGGTTGAGGAAATATCTGCACCTTGAAGCTTCAAGAGAACTTGTCTGCCCTTACGGG 1192
QY 1028 CTGACATCTACAAACCTTCAAGCTGGTTCACTCAAAAAGTGCACAGATCTCAACCTTCTAA 1087
Db 1193 CTGACTCTACAAACCTTAAAGCTGGTTCGCAATTAGTACCTTCAACGCTCACCTCCAG 1252
QY 1088 TACTTAGTGGCTTGGACCTAGTGTGTAATATGAAATCTCTACAGAAATGCAATTTTGG 1147
Db 1253 CCTCGCCCAATTCCAACTCAGTGCCTAATATGTTGCTCTACAAGAAATGGAATTTACT 1312
QY 1148 TCCTCACTACACCAACGACACAGCATCATATATCGATTGAGGGGACGGCTCAG 1207
Db 1313 CTCACATTTGAATCTGAATGCAAAACAGTGTGATCTATGTGACTCGAGGACAAAGGAAG 1372
QY 1208 TGCAAGTCTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCAAG 1267
Db 1373 TTAGAGTTGTAATCTGCAAGGAAATGCCAGTGTTCGACGGTGAAGCTTAGGAGGGGACAAAT 1432
QY 1268 TGCTTGTGGTCCACAGAACTTCGCGTCTGCTGGAAAGTCCCAAGCGGAGAACTTCGAAT 1327
Db 1433 TGCTGTGTGTACCAAGAACTTCGTGTGGCGGAGCAAGCCGAGAAACAGGATTCGAAT 1492
QY 1328 ACCTGGCATTCAGACAGAC 1347
Db 1493 ACATAGTATTCAAGACACAC 1512

RESULT 15

US-09-758-652-14
; Sequence 14, Application US/09758652
; Patent No. 6703544
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-758-652-14
Query Match 10.1%; Score 153.2; DB 4; Length 1689;
Best Local Similarity 49.7%; Pred. No. 3.2e-35;
Matches 726; Conservative 0; Mismatches 578; Indels 156; Gaps 7;
QY 14 AGGAGAACCGGTTCAGATTCCAGCGCTCAATTCGCGAGAGACCTGACATTCGATTGAAT 73
Db 83 AGCTCAACGAGTGCACAACTCAACACCTCAACCGCTTGGAAACCCGACACCCCGTTGAGT 142
QY 74 CAGAGGGCGTTTACATTGAGACTTGGAAACCCCAAACCCAGGAGTTCGAAATGCGCGGG 133
Db 143 TCGAAGGTGGTTTGATTCAAACATGAACTCTCAACACCTGAGCTGAAATGCGCGGGT 202
QY 134 TCGCCCTCTCTCGCTTAGTCTCCCGCGCAACGCCCTTCGTAGGCTTTTCTACTCCAATG 193
Db 203 TCACTGTTTCCAAACTCACCCCTCAACCGCAATGGCTCCACTTGCATCTTACTCACCTT 262
QY 194 CTCGCCAGGAGATCTTCATCCAGCAAGGAAGGGGATACCTTTGGTGTGATATTCCCTGGTT 253
Db 263 ATCCCGGATGATCATCATCGCCCAAGGGAAGAGGACACTGCAGTGCA---AGCCAGGAT 319
QY 254 GTCCTAGACACTATGAAGAGCTCTACACACAGGTCGTCTGATCTCAGTCCCAAAGACCA 313
Db 320 GTCCTGAGACGTTTGAAGGCCCAAGAAACA-----TCAAACAGAAAGGCTC 368
QY 314 CAAGACGTCTCAAGGAGAAGACCAAAGCCAAACAGACAGATAGTCCAGAAAGGTGC 373
Db 369 AAGGTCTG-----CAGAAGCAGCAGCTACAGGACAGTCAACGAAAGATTTC 412
QY 374 ACCGTTTCAGTACGAGGTGATCTCATTCAGTTCACCGGTGTTCTTCTGGCTCTACA 433
Db 413 GTCACCTTCAATGAAGGAGACGTAACCTTCTCTCTGTTCTTCTTACTTGACCTATA 472
QY 434 AGCACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAACAAACAAACCAACC 493
Db 473 ACACCTGGCGATGAACCAAGTGTGTCATCAGTCTTCTTGACACCTTAACCTCAATAACC 532
QY 494 AGCTTGATCAGTTCCCGAGGAGATTCAAATTTGGCTGGGAAACGAGGCAAGAGTTCTTAA 553
Db 533 AGCTTGATCAAAACCCCTAGGGTATTTTACCTTGTCTGGGAACCCAGATATAGAGTACCCAG 592
QY 554 GGTACAGCAACAAGCAGACA----- 575
Db 593 AGACCATGCAACAACAACAACAGCAGCAAGAAAGTCATGGTGGACGCAAGCGGGCAACACC 652
QY 576 -----AAGCAGACGAAGAAGCTTACCATATAGCCCATACAGCCCGCAAGTCAGCTAG 629
Db 653 AGCAGGAGGAAGGAAGAGGTGGCGGCTCTAGTGGCTTCAGCAAAACACTTCTTGG 712
QY 630 ACAAGAAGAGCGTGAATTTAGCCCTTCGAGGACAGCACAGCCCGCAGAGAACGAGCAGGACA 689
Db 713 CACAATCTCTTCAACACCAACGAGGACATAGCTGAGAAACTTTCAGTCTCCAGACGACGAAA 772
QY 690 AGAAGAGAAACGAGGTGGAAACATCT-TCAGCGGCTTCACCGCGGAGTTCTTGGAC 748
Db 773 GGAAGCAGATCGTGACAGTGGAGGAGGTCTCAGCGTTATCAGCCCTATCAGCCCAAGTGGCAAGAAC 832
QY 749 AAGCCCTTCAGGTTGACGACACACAGATAGTGCAAAA----- 785
Db 833 AACAGATGACATGAAGATGAAGAGGAAGATGATGAAGATGAACAAATTCCTCTCACC 892
QY 786 -----CCTAAGAGCGGAGACCGAGAGGTGAAGAAAGAGGGAGGCCATTGTGACAGTGAGGG 938

```
Db      893  CTCCTCCGCGACCAAGCCATGGAAGCGTGAAACAAGACGAGGACGAGACGAGATGAAG 952
Qy      839  GAGGCCTCAGAACTTTGAGCCCGATAGAAAGAGACGTCGCCGACGAAGAAGAGGAATACG 898
Db      953  ATAAACCTCGTCCTAGTCGACCAAGCCAAAGGAAAGCGTGAAACAAGACCAGGACCGAGCG 1012
Qy      899  ATGAAGATGAATATGATACGATGAAGAGGATGAAGGCGTGGCAGG----- 945
Db      1013  AGGACGAAGATGAAGATGAAGATCAACCTCGCAAGAGCGCGCAATGGAGATCGAAAAAGA 1072
Qy      946  -----GGAAAGCAGAGGCGAGGGGAATG 967
Db      1073  CACAACCCAGAGACCTAGACAAGAAGAACACGCTGAAGAGGATGCGAGACAAGAAGCG 1132
Qy      968  GTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAAGAAACATTGGTAGAAACAGATCCC 1027
Db      1133  GGGTTGAGGAAATATCTGCACCTTTGAAGCTTACAGAGAACATTGCTCGCCCTTCACGCG 1192
Qy      1028  CTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACTGCCAAGATCTCAACCTTCTAA 1087
Db      1193  CTGACTTCTACAACCCCTAAAGCTGGTCGCATTAGTACCCCTCAACAGGCTCACCCCTCCAG 1252
Qy      1088  TACTTAGTGCTTGGACCTAGTCTGAATATGAAATCTCTACAGGAATGCAATTGTTG 1147
Db      1253  CCTCCGCGCAATTCCAACTCAGTGCCCAATATGTTGCTCTACAAGATGGAAATTTACT 1312
Qy      1148  TCGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGGGCTCACG 1207
Db      1313  CTCCACATTTGGATCTGAATGCAAAACAGTGTGTATGTGACTCGAGGACAAAGGAAAGG 1372
Qy      1208  TGCAGTCTGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACG 1267
Db      1373  TTAGAGTTGTGAACCTGCCAAGGGAATCAGTGTTCGACGGGTGAGCTTAGGAGGGGACAAT 1432
Qy      1268  TGCTTGTGGTGCCACAGAACTTCGCCGTCGCTGGAAGTCCACAGAGCGAGAACTTCGAAT 1327
Db      1433  TGTGTGTGTACCAACAGAACTTCGTGTGGCGGAGCAAGCGGAGAACAGGATTCGAAT 1492
Qy      1328  ACGTGGCAATTCAGACAGAC 1347
Db      1493  ACATAGTATTCAAGACACAC 1512
```

Search completed: August 24, 2005, 03:57:19
Job time : 267.731 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2005, 00:57:31 ; Search time 921.248 Seconds
(without alignments)
10747.710 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 1524
Sequence: 1 CGGAGCAACCGAGGAGGAGAA.....ctccgaggctgtggttaa 1524

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7316285 seqs, 3248459403 residues

Total number of hits satisfying chosen parameters: 14632570

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_5/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1524	100.0	1524	10	US-09-731-375A-3
2	1524	100.0	1524	14	US-10-728-806-5
3	1524	100.0	1524	17	US-10-100-303A-89
4	1524	100.0	1524	20	US-10-728-323-3
5	1524	100.0	1524	20	US-10-728-051-3
6	1524	100.0	1524	21	US-10-899-551-5
7	1391.2	91.3	1853	21	US-10-958-324-4

Query Match	100.0%	Score 1524;	DB 10;	Length 1524;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1524;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	CGGAGCAACCGAGGAGGAGAA	CGGTCGAGTCCAGCGCTCAATCGCAGAGACTGAC	60
Db	1	CGGAGCAACCGAGGAGGAGAA	CGGTCGAGTCCAGCGCTCAATCGCAGAGACTGAC	60
Qy	61	AATCGATTGATCAGAGCGGCTTACATTGAGACTTGAACCCCAACACGAGGATTC	120	

ALIGNMENTS

RESULT 1
US-09-731-375A-3
; Sequence 3, Application US/09731375A
; Publication No. US20030035810A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0100
; CURRENT APPLICATION NUMBER: US/09/731,375A
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-09-731-375A-3

Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 7280, Ap
Sequence 8531, Ap
Sequence 8592, Ap
Sequence 7583, Ap
Sequence 28860, A
Sequence 7594, Ap
Sequence 2, Appl
Sequence 16, Appl
Sequence 10, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 7674, Ap
Sequence 8556, Ap
Sequence 7642, Ap
Sequence 7670, Ap
Sequence 7613, Ap
Sequence 8587, Ap
Sequence 7603, Ap
Sequence 8164, Ap
Sequence 8133, Ap
Sequence 28863, A
Sequence 14, Appl
Sequence 8584, Ap
Sequence 15, Appl
Sequence 15, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl

Db 61 AATCGCATTAATCAGAGGCGGTACTAATTGAGACTTGGAAACCCCAACAAACAGGAGTTC 120
QY 121 GAATGCGCGCGCTCGCCCTCTCGCTTAGTTCCTCGCGCAACGCCCTTGGTAGGCCT 180
Db 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTTCCTCGCGCAACGCCCTTGGTAGGCCT 180
QY 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCAGCAAGAAAGGGGATACTTTGGGTTG 240
Db 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCAGCAAGAAAGGGGATACTTTGGGTTG 240
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
QY 301 TCCCAAGACCAACAAGAGCTCTCCAGGAGAAAGCAAGCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACAAGAGCTCTCCAGGAGAAAGCAAGCAACAGCAACGAGATAGT 360
QY 361 CACCAGAAGGTGACCGTTCGATGAGGATCTCATTTGCGAGTTCCACCGGTGTGCT 420
Db 361 CACCAGAAGGTGACCGTTCGATGAGGATCTCATTTGCGAGTTCCACCGGTGTGCT 420
QY 421 TTCTGGCTCTACAAACGACGACGACGATGTTGCTGCTTCTTACTGACACCAAC 480
Db 421 TTCTGGCTCTACAAACGACGACGACGATGTTGCTGCTTCTTACTGACACCAAC 480
QY 481 AACAAACGACACGCTTGATCAGTTCCCGCAGGAGATTCAAATTTGGTGGGAACCGAG 540
Db 481 AACAAACGACACGCTTGATCAGTTCCCGCAGGAGATTCAAATTTGGTGGGAACCGAG 540
QY 541 CAAGAGTTCCTTAAGGTACCAACGACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
Db 541 CAAGAGTTCCTTAAGGTACCAACGACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
QY 601 AGCCCTACAGCCGCGAAAGTCTAGCTAGCAGAGAGAGCGTGAATTTAGCCCTCAGGA 660
Db 601 AGCCCTACAGCCGCGAAAGTCTAGCTAGCAGAGAGAGCGTGAATTTAGCCCTCAGGA 660
QY 661 CAGCAGAGCCGACAGAAACGAGCAGACGACAAAGAAAGAAACCAAGTGGAAACATCTTC 720
Db 661 CAGCAGAGCCGACAGAAACGAGCAGACGACAAAGAAAGAAACCAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTACGCGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGACGACAGATAGTG 780
Db 721 AGCGGCTTACGCGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGACGACAGATAGTG 780
QY 781 CAAAACTTAAGCGGACGAGAGCGAGAGTGAAGAGAGGAGCCATTTGTACAGTGGGGGA 840
Db 781 CAAAACTTAAGCGGACGAGAGCGAGAGTGAAGAGAGGAGCCATTTGTACAGTGGGGGA 840
QY 841 GGCTCAGAAATCTTGAGCCGACAGATAGAAAGAGACGTCGCCCAAGAAAGAGGAATACGAT 900
Db 841 GGCTCAGAAATCTTGAGCCGACAGATAGAAAGAGACGTCGCCCAAGAAAGAGGAATACGAT 900
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAAGCAGAGCAGG 960
QY 961 GGGAAATGGTATGAAGAGAGCATCTGCACCGCAAGTGTCTAAAAGAACATTTGGTAGAAC 1020
Db 961 GGGAAATGGTATGAAGAGAGCATCTGCACCGCAAGTGTCTAAAAGAACATTTGGTAGAAC 1020
QY 1021 AGATCCCTCAGCATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCAACAGATCTCAAC 1080
Db 1021 AGATCCCTCAGCATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCAACAGATCTCAAC 1080
QY 1081 CTTCTAATACTTAGGTGGCTTGAACCTTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACTTAGGTGGCTTGAACCTTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTGTGCTCACTACAAACCAAGCAGACGACGATCATATATCGATTGAGGGGACGG 1200
Db 1141 TTGTTGTGCTCACTACAAACCAAGCAGACGACGATCATATATCGATTGAGGGGACGG 1200

QY 1201 GCTCAGTCAAGTCTGTGACACGAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
Db 1201 GCTCAGTCAAGTCTGTGACACGAACCGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
QY 1261 GGTCAAGTCTGTGTGTGTCACAGAACTTTCGCGCTCGTGGAAAGTCCAGAGCGAGAAC 1320
Db 1261 GGTCAAGTCTGTGTGTGTCACAGAACTTTCGCGCTCGTGGAAAGTCCAGAGCGAGAAC 1320
QY 1321 TTGCAATAGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
Db 1321 TTGCAATAGTGGCATTTCAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
QY 1381 AACTCCGTCTATAGATAAACCCTCGCGAGGAGTGGTTGCAAAATTCATATGGCCCTCAAAGG 1440
Db 1381 AACTCCGTCTATAGATAAACCCTCGCGAGGAGTGGTTGCAAAATTCATATGGCCCTCAAAGG 1440
QY 1441 GAGCAGGCAAGGAGCTTAAGAAACAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Db 1441 GAGCAGGCAAGGAGCTTAAGAAACAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGCTGTGGCTTAA 1524

RESULT 2

US-10-228-806-5
; Sequence 5, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; FILE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-228-806-5

Query Match 100.0%; Score 1524; DB 14; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCAGCAACCGGAGGAGAACCGTGCACAGTTCACAGCGCTCAATGCGCAGAGACTGAC 60
Db 1 CGGCAGCAACCGGAGGAGAACCGTGCACAGTTCACAGCGCTCAATGCGCAGAGACTGAC 60
QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACAGGAGTTC 120
Db 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACAGGAGTTC 120
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
QY 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGATACATTTGGGTTG 240
Db 181 TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGATACATTTGGGTTG 240
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
QY 301 TCCCAAGACCAACAGAGAGCTCTCCAAAGAGAGAACCAAGCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACAGAGAGCTCTCCAAAGAGAGAACCAAGCAACAGCAACGAGATAGT 360

Qy 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCCACCGGTGTTGCT 420
Db 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCCACCGGTGTTGCT 420
Qy 421 TTCTGGGTCTACAAACGACGACGACTGATGTTGTTGCTGTTCTCTTTACTGACACCAAC 480
Db 421 TTCTGGGTCTACAAACGACGACGACTGATGTTGTTGCTGTTCTCTTTACTGACACCAAC 480
Qy 481 AACAAACGACAAACAGCTTGATCAGTTCCCAAGGAGATTCAATTTGGCTGGGAAACACGGAG 540
Db 481 AACAAACGACAAACAGCTTGATCAGTTCCCAAGGAGATTCAATTTGGCTGGGAAACACGGAG 540
Qy 541 CAAGAGTCTTTAAGGTACCAACGACCAACAAAGCAGACAAAGCAGACCAAGCTTACCATAT 600
Db 541 CAAGAGTCTTTAAGGTACCAACGACCAACAAAGCAGACAAAGCAGACCAAGCTTACCATAT 600
Qy 601 AGCCCATACGCCCGCAAAAGTCAGCTTAGACAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 601 AGCCCATACGCCCGCAAAAGTCAGCTTAGACAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Qy 661 CAGCACGCGCAGACGAGAAACGAGGAGCAAGAAAGAGAAACGAAAGTGGAAACATCTTC 720
Db 661 CAGCACGCGCAGACGAGAAACGAGGAGCAAGAAAGAGAAACGAAAGTGGAAACATCTTC 720
Qy 721 AGCGGCTTCAGCGCGGAGTTCTTGGAAACAGCTTCCAGGTTGACGACAGACATAGTG 780
Db 721 AGCGGCTTCAGCGCGGAGTTCTTGGAAACAGCTTCCAGGTTGACGACAGACATAGTG 780
Qy 781 CAAACCTTAAGAGCGGAGACGAGAGTGAAGAGAGGAGGACGATTTGACAGTCAAGGGGA 840
Db 781 CAAACCTTAAGAGCGGAGACGAGAGTGAAGAGAGGAGGACGATTTGACAGTCAAGGGGA 840
Qy 841 GGCCTCAGAAATCTTTAGCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAAATACGAT 900
Db 841 GGCCTCAGAAATCTTTAGCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAAATACGAT 900
Qy 901 GAAGATGAATATGAATACGATGAAGAGATAGAGGCGTGCCAGGGAAGCAGAGGAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGATAGAGGCGTGCCAGGGAAGCAGAGGAGG 960
Qy 961 GGGAAATGGTATTGAAGAGAGCAGCTGCACCGCAAGTGTCTAAAGAGACATTTGGTAGAAC 1020
Db 961 GGGAAATGGTATTGAAGAGAGCAGCTGCACCGCAAGTGTCTAAAGAGACATTTGGTAGAAC 1020
Qy 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAAC 1080
Db 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAAC 1080
Qy 1081 CTTCTAATACTAGTGGCTTGGACCTAGTGCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACTAGTGGCTTGGACCTAGTGCTGAATATGGAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTTGTGCTCACTACAAACCAACGACACACAGCATCATATATCGATTGAGGGGACGG 1200
Db 1141 TTGTTTGTGCTCACTACAAACCAACGACACACAGCATCATATATCGATTGAGGGGACGG 1200
Qy 1201 GCTCAGCTGCAAGTTCGTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGCTGCAAGTTCGTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTACAGCTGTTGTTGGTGCACAGAACTTCGCCGTGCTGGAAAGTCCGACGAGGAGAAC 1320
Db 1261 GGTACAGCTGTTGTTGGTGCACAGAACTTCGCCGTGCTGGAAAGTCCGACGAGGAGAAC 1320
Qy 1321 TTCCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGCAGCATAGCCAACTCGCCGGTGAA 1380
Db 1321 TTCCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGCAGCATAGCCAACTCGCCGGTGAA 1380
Qy 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGG 1440
Db 1381 AACTCCGTCTATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGG 1440
Qy 1441 GAGCAGGCAAGGAGCTTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500

Db 1441 GAGCAGGCAAGGAGCTTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Qy 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
RESULT 3
US-10-100-303A-89
; Sequence 89, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-89
Query Match 100.0%; Score 1524; DB 17; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGAGCAACCGGAGGAGAACCGTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTTGAC 60
Db 1 CGGAGCAACCGGAGGAGAACCGTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTTGAC 60
Qy 61 AATCGCATTTGAATCAGAGGGCGGTTACATTTGAGACTTGGAAACCCCAACACCGAGGATTC 120
Db 61 AATCGCATTTGAATCAGAGGGCGGTTACATTTGAGACTTGGAAACCCCAACACCGAGGATTC 120
Qy 121 GAATGGCGCGCGCTCGCCCTCTCTCGTTAGTCTCTCGCGCGCAACCGCCCTTCGTAGGCCT 180
Db 121 GAATGGCGCGCGCTCGCCCTCTCTCGTTAGTCTCTCGCGCGCAACCGCCCTTCGTAGGCCT 180
Qy 181 TTCTACTCCAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGGTTG 240
Db 181 TTCTACTCCAATGCTCCCAAGGAGATCTTCATCAGCAAGAAAGGGGATCTTTGGGTTG 240
Qy 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCGCTCACACAAAGTTCGTGATCTCAG 300
Db 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCGCTCACACAAAGTTCGTGATCTCAG 300
Qy 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
Qy 361 CACCAAGAGGTGCAACCGTTTCGATGAGGGTGATCTCATTCAGGTTCCCAACCGGTTGCT 420
Db 361 CACCAAGAGGTGCAACCGTTTCGATGAGGGTGATCTCATTCAGGTTCCCAACCGGTTGCT 420
Qy 421 TTCTGGCTCTACACGACCAACGACGACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 421 TTCTGGCTCTACACGACCAACGACGACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Qy 481 AACAAACGACAAACCAAGCTTGATTCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAAACACGGAG 540
Db 481 AACAAACGACAAACCAAGCTTGATTCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAAACACGGAG 540
Qy 541 CAAGAGTCTTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db 541 CAAGAGTCTTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Qy 601 AGCCCATACGCCCGCAAAAGTTCAGCTTAGACAAGAGCGTGAAATTTAGCCCTCGAGGA 660

```
Db 601 AGCCCATACAGCCCGCAAAAGTACGCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
Qy 661 CAGCAGACCGCAGAGAAACGAGCAGAGCAAGAAAGAAACCAAGAGTGAAGAAATCTTC 720
Db 661 CAGCAGACCGCAGAGAAACGAGCAGAGCAAGAAAGAAACCAAGAGTGAAGAAATCTTC 720
Qy 721 AGCGGCTTACGCGCGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 AGCGGCTTACGCGCGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Qy 781 CAAAACCTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
Db 781 CAAAACCTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
Qy 841 GGCTCAGAAATCTTGAAGCCAGATAGAAAGAGAGAGCTGCCAGCAAGAAAGAGGAATACGAT 900
Db 841 GGCTCAGAAATCTTGAAGCCAGATAGAAAGAGAGAGCTGCCAGCAAGAAAGAGGAATACGAT 900
Qy 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAAGCAGAGGCAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAAGCAGAGGCAGG 960
Qy 961 GGGAAATGTTAAGAGAGAGATCTGCACCGCAAGTGTCTAAAAGAAACATTTGTTAGAAAC 1020
Db 961 GGGAAATGTTAAGAGAGAGATCTGCACCGCAAGTGTCTAAAAGAAACATTTGTTAGAAAC 1020
Qy 1021 AGATCCCTCAGACATCTCAACCCCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAAC 1080
Db 1021 AGATCCCTCAGACATCTCAACCCCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAAC 1080
Qy 1081 CTTCTAATACCTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACCTTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTGTCTCCTCATACACACCAAGCGCACACAGCATCATATATCGATTGAGGGAGCG 1200
Db 1141 TTGTTGTCTCCTCATACACACCAAGCGCACACAGCATCATATATCGATTGAGGGAGCG 1200
Qy 1201 GCTCAGCTGAAGTCTGTGGAGCAGCAACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGCTGAAGTCTGTGGAGCAGCAACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTCAAGTGTCTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCAAGAGCGAGAAC 1320
Db 1261 GGTCAAGTGTCTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCAAGAGCGAGAAC 1320
Qy 1321 TTGGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380
Db 1321 TTGGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA 1380
Qy 1381 AACTCCGTATAGATTAACCTGCGGAGGAGTGTGTCGAATTCATATGSCCTCCAAAGG 1440
Db 1381 AACTCCGTATAGATTAACCTGCGGAGGAGTGTGTCGAATTCATATGSCCTCCAAAGG 1440
Qy 1441 GAGCAGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACGCTCTCAG 1500
Db 1441 GAGCAGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACGCTCTCAG 1500
Qy 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
```

RESULT 4

```
US-10-728-323-3
; Sequence 3, Application US/10728323
; Publication No. US20040208894A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sogin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
```

```
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0232
; CURRENT APPLICATION NUMBER: US/10/728,323
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
; US-10-728-323-3
```

```
Query Match 100.0%; Score 1524; DB 20; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCAGCAACCGGAGGAGAACGCTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTTGAC 60
Db 1 CGGCAGCAACCGGAGGAGAACGCTGCCAGTTCAGCGCCTCAATGCGCAGAGACCTTGAC 60
Qy 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTTGGAACCCCAACCAACGAGGATTC 120
Db 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTTGGAACCCCAACCAACGAGGATTC 120
Qy 121 GAATGGCGGGCGTGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTGTAGGCCT 180
Db 121 GAATGGCGGGCGTGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTGTAGGCCT 180
Qy 181 TTCTACTCTCAATGCTCCCGCAGGAGATCTTCATCCAGCAAGGAGGGATACCTTTGGGTTG 240
Db 181 TTCTACTCTCAATGCTCCCGCAGGAGATCTTCATCCAGCAAGGAGGGATACCTTTGGGTTG 240
Qy 241 ATATTCCTCGTGTGCTCTAGACACTATGAAGAGCCTCACACAAAGGTCGTGATCTCAG 300
Db 241 ATATTCCTCGTGTGCTCTAGACACTATGAAGAGCCTCACACAAAGGTCGTGATCTCAG 300
Qy 301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGAACCAAGCCCAACAGCAACGAGATAGT 360
Db 301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGAACCAAGCCCAACAGCAACGAGATAGT 360
Qy 361 CACCAGAGGTGCACCGTTTCGATGAGGAGTGTCTCATTTGCAGATTCCTCCACCGGTGTGCT 420
Db 361 CACCAGAGGTGCACCGTTTCGATGAGGAGTGTCTCATTTGCAGATTCCTCCACCGGTGTGCT 420
Qy 421 TTCTGGCTCTCAACGACCAACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAAC 480
Db 421 TTCTGGCTCTCAACGACCAACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACCAAC 480
Qy 481 AACACGCAACACGAGTTGATTCAGTTCCCGAGAGATTCAAATTTGGCTCGGAAACACGGAG 540
Db 481 AACACGCAACACGAGTTGATTCAGTTCCCGAGAGATTCAAATTTGGCTCGGAAACACGGAG 540
Qy 541 CAAGAGTCTTAAAGGTACCAGCAACAAAGCAGCAAAAGCAGACGAAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAAGGTACCAGCAACAAAGCAGCAAAAGCAGACGAAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAAAGTCAAGCTTAGAAGAGAGCGTGAAATTTTAGCCCTCGAGGA 660
Db 601 AGCCCATACAGCCCGCAAAAGTCAAGCTTAGAAGAGAGCGTGAAATTTTAGCCCTCGAGGA 660
Qy 661 CAGCAGACCGCAGAGAAACGAGCAGAGCAAGAAAGAAACCAAGAGTGAAGAAATCTTC 720
Db 661 CAGCAGACCGCAGAGAAACGAGCAGAGCAAGAAAGAAACCAAGAGTGAAGAAATCTTC 720
Qy 721 AGCGGCTTCAAGCGCGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 721 AGCGGCTTCAAGCGCGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGTG 780
Qy 781 CAAAACCTAAGAGGGCAGAGACCGAGAGTGAAGAGGAGGCCATTTGTGACAGTGAAGGGA 840
```

Db 781 CAAAACCTAAGAGCGGAGACGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGAGGGA 840
Qy 841 GGCTCTAGAAATCTTGAACCCAGATAGAAAGAGAGCGTCCGACGAAAGAGAGGAATACGAT 900
Db 841 GGCTCTAGAAATCTTGAACCCAGATAGAAAGAGAGCGTCCGACGAAAGAGAGGAATACGAT 900
Qy 901 GAAGATGAATATGAATACGATGAAGAGAGATAGAAAGGCGTGGCAGCGGAAGCAGAGGAGG 960
Db 901 GAAGATGAATATGAATACGATGAAGAGAGATAGAAAGGCGTGGCAGCGGAAGCAGAGGAGG 960
Qy 961 GGGAAATGGTATGAAGAGAGCGATCTGCAACGCGAAGTCTTAAAGAGACATTTGGTAGAAAC 1020
Db 961 GGGAAATGGTATGAAGAGAGCGATCTGCAACGCGAAGTCTTAAAGAGACATTTGGTAGAAAC 1020
Qy 1021 AGATCCCTGACATCTTCAACCCCTCAAGCTGGTTCACTCAAACTGCGCAACGATCTCAAC 1080
Db 1021 AGATCCCTGACATCTTCAACCCCTCAAGCTGGTTCACTCAAACTGCGCAACGATCTCAAC 1080
Qy 1081 CTTCTAATACCTTAGTGGCTTGAACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 1081 CTTCTAATACCTTAGTGGCTTGAACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Qy 1141 TTGTTTGTGCTCACTACAACCAACGCGACACAGCATATATATCGATTGAGGGACGG 1200
Db 1141 TTGTTTGTGCTCACTACAACCAACGCGACACAGCATATATATCGATTGAGGGACGG 1200
Qy 1201 GCTCAGTGAAGTCTGAGAGCAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 1201 GCTCAGTGAAGTCTGAGAGCAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Qy 1261 GGTCAAGTGTGTTGTCACAGAACTTCGCGTCTGCGAAGTCCGAGAGCGAGAAC 1320
Db 1261 GGTCAAGTGTGTTGTCACAGAACTTCGCGTCTGCGAAGTCCGAGAGCGAGAAC 1320
Qy 1321 TTGGAATACGTGTCATTAAGACAGACTCAAGGCCAGCATAGCAACCTCGCGGTGAA 1380
Db 1321 TTGGAATACGTGTCATTAAGACAGACTCAAGGCCAGCATAGCAACCTCGCGGTGAA 1380
Qy 1381 AACTCGTCTATAGATAACCTGCGGAGGAGGTGTTGCAAAATCATATGCGCTCCAAAGG 1440
Db 1381 AACTCGTCTATAGATAACCTGCGGAGGAGGTGTTGCAAAATCATATGCGCTCCAAAGG 1440
Qy 1441 GAGCAGCAAGGAGCTTAAGAAACAAACCCCTTCAGTTCTTGGTCCACCGTCTCAG 1500
Db 1441 GAGCAGCAAGGAGCTTAAGAAACAAACCCCTTCAGTTCTTGGTCCACCGTCTCAG 1500
Qy 1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524
Db 1501 CAGTCTCCGAGGCGTGTGGCTTAA 1524

RESULT 5
US-10-728-051-3
; Sequence 3, Application US/10728051
; Publication No. US20040234548A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Bottomly H., Kim
; APPLICANT: Sosin B., Howard
; APPLICANT: Burks A., Wesley
; APPLICANT: Sampson A., Hugh
; TITLE OF INVENTION: Microbial Delivery System
; FILE REFERENCE: 2002834-0222
; CURRENT APPLICATION NUMBER: US/10728,051
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: 60/195,035
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 09/731,375
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3

; LENGTH: 1524
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-728-051-3
Query Match 100.0%; Score 1524; DB 20; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCACAAACCGGAGGAGAAACGCGTCCAGTTCCAGCGCTCAATGCGAGAGACCTGAC 60
Db 1 CGGCACAAACCGGAGGAGAAACGCGTCCAGTTCCAGCGCTCAATGCGAGAGACCTGAC 60
Qy 61 AATCGCATTAATCAGAGGGCGTTACATTGAGACTTTGGAACCCCAACAAACGAGATTTC 120
Db 61 AATCGCATTAATCAGAGGGCGTTACATTGAGACTTTGGAACCCCAACAAACGAGATTTC 120
Qy 121 GAATGCGCGGGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 121 GAATGCGCGGGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Qy 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCAATCCAGAAAGGAGGATACTTTGGGTTG 240
Db 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCAATCCAGAAAGGAGGATACTTTGGGTTG 240
Qy 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCAACAAGGTGCTGATCTTCAG 300
Db 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCTTCAACAAGGTGCTGATCTTCAG 300
Qy 301 TCCCAAGAACCCACCAAGAGCTCTCCAAAGGAGAGACCAAGGCCAAGCAACGAGATAGT 360
Db 301 TCCCAAGAACCCACCAAGAGCTCTCCAAAGGAGAGACCAAGGCCAAGCAACGAGATAGT 360
Qy 361 CACCAGAGGTGACCGTTTCCAGATGAGGCTGATCTCATTCAGATTCCCAACCGGTGCT 420
Db 361 CACCAGAGGTGACCGTTTCCAGATGAGGCTGATCTCATTCAGATTCCCAACCGGTGCT 420
Qy 421 TTCTGCTCTCAACAACGACCACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 421 TTCTGCTCTCAACAACGACCACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
Qy 481 AACACGACCAACCGCTTCCAGTCCCGAGGAGATTCATTTGGCTGGGACACGGAG 540
Db 481 AACACGACCAACCGCTTCCAGTCCCGAGGAGATTCATTTGGCTGGGACACGGAG 540
Qy 541 CAAGAGTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 541 CAAGAGTCTTAAGGTACCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Qy 601 AGCCCATACAGCCCGCAAAAGTCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 601 AGCCCATACAGCCCGCAAAAGTCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Qy 661 CAGCAGCGCGCAGAGAACGCGACAGAGAGAGAGAAACGAGGTTGGAACATCTTC 720
Db 661 CAGCAGCGCGCAGAGAACGCGACAGAGAGAGAGAAACGAGGTTGGAACATCTTC 720
Qy 721 AGCGGCTTCAAGCGGAGTTTCTTGGAAACAAAGCTTCCAGGTTGACACAGAGATAGT 780
Db 721 AGCGGCTTCAAGCGGAGTTTCTTGGAAACAAAGCTTCCAGGTTGACACAGAGATAGT 780
Qy 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGGAGGAGCCATTGTGACAGTGGGA 840
Db 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGGAGGAGCCATTGTGACAGTGGGA 840
Qy 841 GGCCTCAGAAATCTTGAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 GGCCTCAGAAATCTTGAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 901 GAAGATGAATATGAATACGATGAAGGAGGATAGAAAGCGGTGGCGAGGAGAGAGAGAG 960
Db 901 GAAGATGAATATGAATACGATGAAGGAGGATAGAAAGCGGTGGCGAGGAGAGAGAGAG 960

QY 961 GGAATGGTATTGAAGAGACGATCTGCA CGCAAGTCTCTAAAGAAACATTTGGTAGAAAC 1020
DB 961 GGGAAATGGTATTGAAGAGACGATCTGCA CGCAAGTCTCTAAAGAAACATTTGGTAGAAAC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
DB 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
QY 1081 CTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA 1140
DB 1081 CTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCA 1140
QY 1141 TTGTTTGTCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
DB 1141 TTGTTTGTCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
QY 1201 GCTCAGTGAAGTCTGAGACGACAA CGGCAACAGAGTGTAACA GAGAGCTTCAAGAG 1260
DB 1201 GCTCAGTGAAGTCTGAGACGACAA CGGCAACAGAGTGTAACA GAGAGCTTCAAGAG 1260
QY 1261 GGTCACTGCTTGTGGTCCACAGAACTTCGCCGCTCGCTGGAAAGTCCAGAGCGAGAAC 1320
DB 1261 GGTCACTGCTTGTGGTCCACAGAACTTCGCCGCTCGCTGGAAAGTCCAGAGCGAGAAC 1320
QY 1321 TTGCAATAGCTGGCAATTCAGAAGACAGACTCAAGGCCAGCATATGCCAACCTCGCCGCTGAA 1380
DB 1321 TTGCAATAGCTGGCAATTCAGAAGACAGACTCAAGGCCAGCATATGCCAACCTCGCCGCTGAA 1380
QY 1381 AACTCGTCTATAGTAATACCTGCCGAGAGAGTGGTTGCAAAATTCATATGGCTTCAAAGG 1440
DB 1381 AACTCGTCTATAGTAATACCTGCCGAGAGAGTGGTTGCAAAATTCATATGGCTTCAAAGG 1440
QY 1441 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCAAGTTCTTCCACCGTCTCAG 1500
DB 1441 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCAAGTTCTTCCACCGTCTCAG 1500
QY 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
DB 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524

RESULT 6
US-10-899-551-5
; Sequence 5, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burke, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899, 551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1524
; TYPE: DNA
; ORGANISM: species Arachis hypogaea
US-10-899-551-5

Query Match 100.0%; Score 1524; DB 21; Length 1524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGACGACACCGGAGGAGACCGTCCAGTTCACAGCGCTCAATTCGCGAGAGACCTGAC 60
DB 1 CGGACGACACCGGAGGAGACCGTCCAGTTCACAGCGCTCAATTCGCGAGAGACCTGAC 60
QY 61 AATCGCATTTGAATCAGAGGCGGTTTACATTTGAACTTGAACCCCAACCGAGAGTTC 120

DB 61 AATCGCATTTGAATCAGAGGCGGTTTACATTTGAACTTGAACCCCAACCGAGAGTTC 120
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTCTCGTAGSCCT 180
DB 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTCTCGTAGSCCT 180
QY 181 TTCTACTCCAAATGCTCCCCAGAGATCTTTCATCCAGCAAGAGGAGGATCTTTTGGGTTG 240
DB 181 TTCTACTCCAAATGCTCCCCAGAGATCTTTCATCCAGCAAGAGGAGGATCTTTTGGGTTG 240
QY 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCAACAAGGTCGTCTGATCTCAG 300
DB 241 ATATTCCCTGGTTGCTCTAGACACTATGAAGAGCTTCAACAAGGTCGTCTGATCTCAG 300
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAAGAGAGACCCAAAGCCAAACAGCAACGAGATAGT 360
DB 301 TCCCAAGACCAACCAAGAGCTCTCCAAAGAGAGACCCAAAGCCAAACAGCAACGAGATAGT 360
QY 361 CACCAGAGGTGCAACCGTTTCGATGAGGTGATCTCAATTCGAGTTCCTTCTGACCAAC 420
DB 361 CACCAGAGGTGCAACCGTTTCGATGAGGTGATCTCAATTCGAGTTCCTTCTGACCAAC 420
QY 421 TTCTGGCTCTCAACGACCAACGACTGATGTTGCTGTTGCTGTTCTTCTGACCAAC 480
DB 421 TTCTGGCTCTCAACGACCAACGACTGATGTTGCTGTTGCTGTTCTTCTGACCAAC 480
QY 481 AACACGACCAACCGCTTGCATGTTCCCGAGAGATTCATTTGCTGGGAAACCGGAG 540
DB 481 AACACGACCAACCGCTTGCATGTTCCCGAGAGATTCATTTGCTGGGAAACCGGAG 540
QY 541 CAAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
DB 541 CAAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
QY 601 AGCCCATACAGCCCGAAAGTCCAGCTAGACAAAGAGAGCTGAAATTTAGCCCTCGAGA 660
DB 601 AGCCCATACAGCCCGAAAGTCCAGCTAGACAAAGAGAGCTGAAATTTAGCCCTCGAGA 660
QY 661 CAGCACAGCCGACAGAGAACGAGCAGACAAAGAGAGAAACGAAAGTGGAAACATCTTC 720
DB 661 CAGCACAGCCGACAGAGAACGAGCAGACAAAGAGAGAAACGAAAGTGGAAACATCTTC 720
QY 721 AGCGGCTTACGCCCGAGTCTCTGGAACAAAGCTTTCAGGTTGACGACAGATAGTG 780
DB 721 AGCGGCTTACGCCCGAGTCTCTGGAACAAAGCTTTCAGGTTGACGACAGATAGTG 780
QY 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGCCATTTGTGACGTGAGGGA 840
DB 781 CAAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGCCATTTGTGACGTGAGGGA 840
QY 841 GGCCTCAGAACTCTGAGCCACAGATAGAAAGACGCTGCCGACGAAAGAGAGAAATACGAT 900
DB 841 GGCCTCAGAACTCTGAGCCACAGATAGAAAGACGCTGCCGACGAAAGAGAGAAATACGAT 900
QY 901 GAAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAGCAGAGCAGG 960
DB 901 GAAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAGCAGAGCAGG 960
QY 961 GGAATGTTTGAAGAGACGATCTGCAACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
DB 961 GGAATGTTTGAAGAGACGATCTGCAACCGCAAGTCTTAAAGAAACATTTGGTAGAAAC 1020
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
DB 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
QY 1081 CTTCTAATCTTAGTGGCTTGGACCTTAGTCTGAATATGGAATCTTCTACAGGAATGCA 1140
DB 1081 CTTCTAATCTTAGTGGCTTGGACCTTAGTCTGAATATGGAATCTTCTACAGGAATGCA 1140
QY 1141 TTGTTTGTCTCACTCAACCAACCGACACAGCATCATATATCGATTGAGGGACGG 1200
DB 1141 TTGTTTGTCTCACTCAACCAACCGACACAGCATCATATATCGATTGAGGGACGG 1200

QY 1201 GCTCAGCTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db |||||
QY 1201 GCTCAGCTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db |||||
QY 1261 GGTCAAGCTGCTTGTGGTGGCCACAGAACTTTCGCGCTGCTGGAAAGTCCCGAGAGCGAAGC 1320
Db |||||
QY 1261 GGTCAAGCTGCTTGTGGTGGCCACAGAACTTTCGCGCTGCTGGAAAGTCCCGAGAGCGAAGC 1320
Db |||||
QY 1321 TTGCAATACGTGGCAATTCAGACAGAGCTCAAGGCCCCAGCATAGCAACCTTCGCGGTGAA 1380
Db |||||
QY 1321 TTGCAATACGTGGCAATTCAGACAGAGCTCAAGGCCCCAGCATAGCAACCTTCGCGGTGAA 1380
Db |||||
QY 1381 AACTCCGCTCATAGTAACTCCCGAGAGGTGGTGGCAATTCATATGGCTCCAAAGG 1440
Db |||||
QY 1381 AACTCCGCTCATAGTAACTCCCGAGAGGTGGTGGCAATTCATATGGCTCCAAAGG 1440
Db |||||
QY 1441 GAGCAGGCAAGGCGAGCTTAAAGAACAAACCCCTTCAAGTTCCTTCGTTCCACCGTCTCAG 1500
Db |||||
QY 1441 GAGCAGGCAAGGCGAGCTTAAAGAACAAACCCCTTCAAGTTCCTTCGTTCCACCGTCTCAG 1500
Db |||||
QY 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db |||||
QY 1501 CAGTCTCCGAGGGCTGTGGCTTAA 1524
Db |||||

RESULT 7
US-10-958-324-4
; Sequence 4, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIQUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; PRIOR FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1853
; TYPE: DNA
; ORGANISM: Arachis hypogaea
US-10-958-324-4

Query Match 91.3%; Score 1391.2; DB 21; Length 1853;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 1454; Conservative 0; Mismatches 68; Indels 2; Gaps 2;

QY 2 GGCAGCAACCGAGGAGAACGCGTGCAGTTCAGCGCCCTCAATGGCGCAGAGACCTTGACA 61
Db |||||
QY 71 GGCAGCAGCCGAGGAGAAATGCGTCCAGTTCAGCGCCTCAATGGCGCAGAGACCTTGACA 130
Db |||||
QY 62 ATGCAATTAATCAGAGGGCGGTACATTCAGACTTGGAAACCCCAACACACAGAGTTGCG 121
Db |||||
QY 131 ACCGCAATTAATCGGAGGGCGGTACATTCAGACTTGGAAACCCCAACACAGAGTTGCG 190
Db |||||
QY 122 AATGCGCGGCGCTCGCCCTCTCGCTTACTGCTCCGCGCAACGCCCTTCGTAGGCCCTT 181
Db |||||
QY 191 AATGCGCGGCGCTCGCCCTCTCGCTTACTGCTCCGCGCAACGCCCTTCGTAGGCCCTT 250
Db |||||
QY 182 TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAGTACTTTGGGTGA 241
Db |||||
QY 251 TCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGAGTACTTTGGGTGA 310
Db |||||
QY 242 TATTCCCTGTTGCTCTAGACACTATGAAGAGCGCTCACACAAAGGTCGTGATCTCAGT 301
Db |||||

Db |||||
QY 311 TATTCCCTGTTGCTCTAGCACATATGAAGAGCCTGCACAAAGGAGCGCCGATATCAGT 370
QY 302 CCCAAAGACCACCAAGACGCTCTCCAAAGGAGAAAGCAACCAAGCCCAACAGCAACGAGATAGTC 361
Db |||||
QY 371 CCCAAAGACCACCAAGACGCTTTCGAAGAGAAAGCAACCAAGCCCAACAGCAACGATAGTC 430
QY 362 ACCAGAAAGTGCAACCGTTTCGATGAGGGTGATCTCATTTGAGTTCCCAACCGTGTGCTT 421
Db |||||
QY 431 ACCAGAAAGTGCAACCGTTTCAATGAGGGTGATCTCATTTGAGTTCCCAACCGTGTGCTT 490
QY 422 TCTGGCTCTACAAAGCAGCAGCACTGATGTTGCTGCTTTCTCTTACTGACCAACA 481
Db |||||
QY 491 TCTGGCTGTACAACAGCAGCAGCACTGATGTTGCTGCTTTCTCTTACTGACCAACA 550
Db |||||
QY 482 ACAACGACAACCAAGCTTGATTCAGTTCCTCCAGAGAGATTCATTTTGGGTGGAAACCGGAGC 541
Db |||||
QY 551 ACAACGACAACCAAGCTTGATTCAGTTCCTCCAGAGAGATTCATTTTGGGTGGAAACCGGAGC 610
QY 542 AAGAGTTCTTAAGGTACCGAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATATA 601
Db |||||
QY 611 AAGAGTTCTTAAGGTACCGAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATATA 670
QY 602 GCCCATACAGCCCGCAAGTTCAGCTTAGACAGAGAGCGTGAATTTAGCCCTCGAGGAC 661
Db |||||
QY 671 GCCCATACAGCCCGCATAGTTCGCGCTTAGACAGAGAGCGTGAATTTGCGCCCTCGAGGAC 730
QY 662 AGCAGAGCCGCGAGAGAACGAGCAGGACAAAGAAAGAAACGAAAGGTGGAACATCTTCA 721
Db |||||
QY 731 AGCAGAGCCGCGAGAGAACGAGCAGGACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 790
QY 722 GGGCTTACGCGCGAGTTTCTGGAAACAAAGCCTTCCAGGTTGACGACAGACAGATAGTC 781
Db |||||
QY 791 GGGCTTACGCGCGAGTTTCTGGAAACAAAGCCTTCCAGGTTGACGACAGACAGATAGTC 850
QY 782 ARAACCTTAGAGCGGAGACGAGAGTGAAGAGAGGAGGCCATTTGACAGAGAGGAGGAG 841
Db |||||
QY 851 ARAATCTGTGGGCGAGAGACGAGAGTGAAGAGAGGAGGCCATTTGACAGGAGGAGGAG 910
QY 842 GCCTCAGAAATCTTGAGCCCGAGATAGAAAGAGAGCTGCCAGCAAGAGAGAGAGATAGCATG 901
Db |||||
QY 911 GCCTCAGAAATCTTGAGCCCGAGATAGAAAGAGAGTGCAGAGAGAGAGAGAGAGAGATAGCATG 970
QY 902 AAGATGAATATGAATACGATGAAGAGGATAGAAAGCGCTGCGAGGGAGAGAGAGAGAGGAG 961
Db |||||
QY 971 AAGATCAATATGAATACCAAGATGGAAGCGGTGCGAGGGGAGAGAGAGAGAGAGAGAGAG 1030
QY 962 GGAATGGTATTGAAGAGAGAGATGCAACCGCAAGTCTGCAAAAGAGAGAGAGAGAGAGAGAG 1021
Db |||||
QY 1031 GGAATGGTATTGAAGAGAGAGATGCAACCGCAAGTCTGCAAAAGAGAGAGAGAGAGAGAGAG 1090
QY 1022 GATCCCTGACATCTACAAACCTCTCA-AGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
Db |||||
QY 1091 GATCCCTGACATCTACAAACCTCTCAAGCTCTGAGCTGGTTCACTCAAACTGCC-ACGATCTCAAC 1149
QY 1081 CTTCTAATCTTAGTGCTTGGACCTAGTGTGATATGGAATCTCTACAGAGATGCA 1140
Db |||||
QY 1150 CTTCTAATCTTAGTGCTTGGACTTAGTGTGATATGGAATCTCTACAGAGATGCA 1209
QY 1141 TTGTTTGTGCTCATTACAAACCAACGACACAGCATCATATATCGATTTAGGGAGCGG 1200
Db |||||
QY 1210 TTGTTTGTGCTCATTACAAACCAACGACACAGCATCATATATGCAITTAGGGAGCGG 1269
QY 1201 GCTCAGTCAAGTGTGGACAGCAACGCAAGAGTGTACGACGAGGAGGCTTCAAGAG 1260
Db |||||
QY 1270 GCTCAGTCAAGTGTGGACAGCAACGCAAGAGTGTACGACGAGAGCTTCAAGAG 1329
QY 1261 GGTTCAGTCTTGTGTGCGCACAGAACTTTCGCGCTGCTGGAAGTCCCGAGAGCGAGAAC 1320
Db |||||
QY 1330 GGTTCAGTCTTGTGTGCGCACAGAACTTTCGCGCTGCTGGAAGTCCCGAGAGCGAGAAC 1389
QY 1321 TTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCAGCATATAGCCAACTCTGCGCGGTGAA 1380
Db |||||

Db 1390 TTCGAATACGTGGCATTTCAAGACAGATTCAAGGCCAGCATGCCAACTTTGCGGTGAA 1449
QY 1381 AACTCCGTATAGATAACCTGCCGAGGAGGTGTTGGCAATTCATATGCGCTCCAAAGG 1440
Db 1450 AACTCTTCATAGATAACCTGCCGAGGAGGTGTTGGCAATTCATATGCGCTCCCAAG 1509
QY 1441 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCCTCCACCGTCTCAG 1500
Db 1510 GAGCAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCCTCCACCGTCTCAG 1569
QY 1501 CAGTCTCCGAGGCTGTGCTTAA 1524
Db 1570 CAGTCTCCGAGGCTGTGCTTAA 1593

RESULT 8

US-09-758-652-13
; Sequence 13, Application US/09758652
; Patent No. US20010011377A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14,1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-758-652-13
Query Match 34.0%; Score 517.4; DB 9; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.7e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;
QY 5 AGCAACGGAGGAACGGGTGCCAGTTCAGGCGCTCAATGCGCAGAGACCTGACAATC 64
Db 71 AGCAGCCACAGCAAAACGAGTGCAGATCCAAAGCCTCAATGCGCTTAAACCGGATAAC 130
QY 65 GCATTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCACACACCGAGGTTCCGAAT 124
Db 131 GTATAGAGTCAGAAGGTGGCTTCATTGAGACATGGAAACCGCAACCAAGCCATTCCAGT 190

QY 125 GGGCGGGGTGCGCCCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCTGTAGGCTTTCT 184
Db 191 GTGCGGGTGTGCGCTCTCTCGCTCACCTCAACGGCAACGCCCTTCTGCGACCTTCT 250
QY 185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATACCTTTGGTGTGATAT 244
Db 251 ACACCAACGCTCCCGAGGAGATCTACACCAAGAGTAGTGGTATTTTGGCATGATAT 310
QY 245 TCCCTGTGTCTTAGACACTATGAAGAGCCTCACACAAAGGTGCTCGATCTCAGTCCC 304
Db 311 TCCCGGGTGTCTTAGCACTTTGAAGACCT----- 342
QY 305 AAAGACCACCAAGAGCTCTCCAAGGAGAGACCAAAAGCCAAACAGCAACGAGATAGTCACC 364
Db 343 -----CAACAAAAGGACAAAGCAGCAGGCCCCCAAGACCGTCACC 382
QY 365 AGAAGGTGACCGTTTCGATGAGGGTGATCTCATTTGCAGTTCCACCGGTGTGCTTTCT 424
Db 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTTGCAGTGCAACCGGTTTGCATCT 442
QY 425 GGCTCTACAAACGACACGACACTGATGTTGTGCTTCTCTTACTGACACCAACAACA 484
Db 443 GGATGTACAAACANTGAAGACACTCTCTGTTGTTGCGTTCCTCTTATTTGACCAACAGCT 502
QY 485 ACGACAACCCAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAAACACGGAGCAAG 544
Db 503 TCCAGAACCGCTCGACAGATGCTCTAGGAGATTCTATCTCTGCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAGCAAGAGCTTACCATATAGCC 604
Db 563 AGTTCTTACAGTATCAGCCACAGAAAGCAGCAAG----- 595
QY 605 CATACAGCCGCAAGTTCAGCCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACAGC 664
Db 596 -----GAGGTACTC 604
QY 665 ACAGCCGAGAGAACGAGCAGGACGAAGAAGAAAGAAAGAGGTGGAAACATCTTCAGCG 724
Db 605 AAAGCCAGAAAGAAAGCGTCCAGCAAGAAAGAAAGAAAGAGGAGGAGCAGCATATTGAGTG 664
QY 725 GCTTACGCGCGAGTTCCTGGAAACAGCTTCCAGGTTGACGACAGACAGATAGTCAAA 784
Db 665 GCTTCCGCGCGGATTTCTTGGAAACATGCGTTC---GTGCTGACAGGCAAGATAGTGAGAA 721
QY 785 ACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGAGGCC 844
Db 722 AGCTCAAGGTGAGAACGAGAGGAGAGAGAGGGTGCCCATTTGTGACAGTGAAGAGGATC 781
QY 845 TCAGAAATCTTGAGCCAGATAGAAAGAGACGTTGCCGACGAAGAAGAGGAAATACGATGAAG 904
Db 782 TCAGCGTGATAAGCCACCCACGGAAGAGCAGCAACAAAGACCCCGAGGAAGAGGAGAACG 841
QY 905 ATGAATATGAATACGATGAAGAGGATAGAGCGTGGCAGGAGGAGCAGAGCAGGGGGA 964
Db 842 CAGATTGTGACAGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
QY 965 ATGGTATTGAAGAGACGATCTGACCGCAAGTGCTAAAGAAACATTGTGTAAGAAACAGAT 1024
Db 887 ATGGCATTTGACGAGACCATTTGCAATGAGACTTCGCCACAAACATTGGCCAGACTTCA 946
QY 1025 CCCCTGACATCTCAACCCCTCAAGCTGGTTCTCACTCAAAAACCTGCCAACGATCTCAACCTTC 1084
Db 947 CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCACAAACCGCTACCAGCCTCGACTTC 1006
QY 1085 TAATTAAGTGGTGGCTTGGACCTAGTGTGATATAGGAATCTCTACAGGATGCAATGT 1144
Db 1007 CAGCCCTCTCGTGGGTCAAAACTCAGTGGCCAGTTTGGATCTACTCCGCAAGATGCTATGT 1066
QY 1145 TTGTGCTCACTACAAACCAACGACAGCATCATATATCGATTTGAGGGGACGGGCTC 1204
Db 1067 TCGTGCCACACTACAAACCTGAAACGCAACAGCATATATACGATTTGANTGAGCGGCAAT 1126

QY 1205 ACGTGAAGTCTGTGACAGCAACGGCAACAGAGTGTACGAGAGGCTTCAAGAGGGTC 1264
DB 1127 TGGTACAGTGTGTAATGCAATGGTGAGAGAGTGTGATGAGAGCTGCAAGAGGGAC 1186
QY 1265 ACGTGTGTGTGCTCCACAGAACTTCCCGCTGCTGGAAAGTCCAGAGCGAGAACTTCG 1324
DB 1187 AGGTGTTAATGTGCCAACAACCTTGGGGTGGCTGCAAGATCAACAGAGCGAACAATTCG 1246
QY 1325 AATACGTGGCAATCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAACT 1384
DB 1247 AGTATGTTTCAATCAAGACCAATGATAGACCTCGATCGGCAACTTGGAGGTGCAAACT 1306
QY 1385 CCGTCATAGATAAACCCTGCCGAGAGGTGGTGGCAAAATTCATATGGCTCCAAAGGGAGC 1444
DB 1307 CATTTGTTGAACGATTGCCGAGAGAGTGTTCAGCAAACTTTTAACTAAGGAGGAGC 1366
QY 1445 AGCAAGCGAGCTTAAGAAACAACCCCTTCAAGTCTTGGTTCCACCTCTCAGCAGT 1504
DB 1367 AGGCCAGCGAGTCAAGAAACAACCTTTTCAAGTCTTGGTTCCACCTTAAGGAGTCTC 1426
QY 1505 CTCCGAGGGCTGTGGCTTA 1523
DB 1427 AGAGGAGGTGTGGCTTA 1445

RESULT 9

US-10-684-651-13
; Sequence 13, Application US/10684651
; Publication No. US20040064858A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/684,651
; FILING DATE: 14-Oct-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-684-651-13

Query Match 34.0%; Score 517.4; DB 18; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.7e-154; Indels 144; Gaps 4;
Matches 954; Conservative 0; Mismatches 421;
QY 5 AGCAACCGGAGGAGAACGCGTGCAGTTCAGAGCCCTCAATGGCGAGAGACTTGACAATC 64
DB 71 AGCAGCCACAGCAAAACGAGTCCAGATCCAAAGCTCAATGCCCTAAACCCGATTAACC 130
QY 65 GCATTGAATCAGAGGGCGGTTCATTGAGACTTTGGAAACCCCAACAACAGAGTTCGAAT 124
DB 131 GTATAGAGTCAGAAAGTGGCTTCATTGAGACATGGAAACCTTAACAACAAGCCATTCCAGT 190
QY 125 GCCCGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCTTTCT 184
DB 191 GTCCGGTGTTCCTCTCTCGCTTGACCTCAACCGCAACGCCCTTCGACAGACTTCCT 250
QY 185 ACTTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTGATAT 244
DB 251 ACACCAACGCTCCAGGAGATCTACATCCAAAGGTAGTGTATTTTGGCATGATAT 310
QY 245 TCCCTGGTGTCTAGACACTATGAAGAGCTCTACACACAAGGTCTCGATCTCAGTCCC 304
DB 311 TCCCGGTTGTCTTAGCACATTTGAAGAGCTT----- 342
QY 305 AAAGACCAACAAGAGCTCTCCAGGAGAGAACCAAGCCAAACAGCAAGAGATAGTCACC 364
DB 343 -----CAACAAAAGGACAAAGCAGCAGGCCCCCAAGACCGTCAACC 382
QY 365 AGAAGGTGACACCGTTTCGATGAGGTGATCTCATTTGCAAGTTCACCGGTGTTGCTTTCT 424
DB 383 AGAAGATCTATCACTTCAGAGAGGGTGTATTGATTGCAAGTCCCAACCGGTTTGCATCT 442
QY 425 GGCTCTCAACAGCACGACACTGTGTTGCTGCTTTCTTTACTGACACCAACAACA 484
DB 443 GGATGTACAACAATGAAGACACTCTCTGTTGTTCCGCTTTCTTTATTTGACACCAACAGCT 502
QY 485 ACAGCAACAGCTTGTATCAGTTCCTCCAGAGATTTCAATTTGGCTGGGAAACCGGAGCAAG 544
DB 503 TCCAGAAACAGCTCGACAGATGCTTAGGAGATTTCTATCTTGTCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCCTTAAGGTACCAGCAACAAAGCAGACAAAAGCAGACGAAGAGCTTACCATATAGCC 604
DB 563 AGTTCTACAGTATCAGCCACAGACGACGACG- 595
QY 605 CATACAGCCGCAAAAGTTCAGCTTAGACAAAGAGCGTGAATTTAGCCCTTCGAGGACAGC 664
DB 596 -----GAGGTACTC 604
QY 665 ACAGCCGAGAGAAACGAGCAGGACAAAGAAAGAAACGAGAGTGGAAACATCTTCAGCG 724
DB 605 AAAGCCAGAAAGGAAAGCGCTCAGCAAGAAAGAAAGAAACGAAAGGAGCGCATATTGAGTG 664
QY 725 GCTTTCAGCGCGGAGTTCTCTGGAAACAAGCCCTTCCAGGTTCAGCAGACAGATAGTGAAA 784
DB 665 GCTTCGCGCGGAAATTCCTGGAAACATGCGCTTC---GTCTGGGACAGGAGATAGTGA 721
QY 785 ACCTAAGAGCGCAGACCGGAGGTGAAGAAAGAGGAGCCATTTGTGACAGTGGGGAGGCC 844
DB 722 AGCTACAAGGTGAGAAACGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
QY 845 TCAGAAATCTTGAGCCCGCAGATAGAAAGAGAGCTCCGACGAGAAAGAGGAAATAGATGAAG 904
DB 782 TCAGCGTGTATAGCCCAACCCCAAGAGAGCAGCAACAAGAGCCCGAGGAGAGGAGGAGGAGC 841
QY 905 ATGAATATGAATACGATAGAGGATAGAGGCGTGGCGGGGAGAGCAGAGGAGGAGGAGGAG 964
DB 842 CAGATTGTGACGAGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 886
QY 965 ATGGTATTGAAGAGAGGAGTCTGCACCGCAAGTCTTAAAGAAACATTTGGTAGAAACAGAT 1024
DB 887 ATGGCATTGACGAGACCAATTTGCACATGAGACTTCGCCACACACATTTGGCCACACTTCAT 946

QY 1025 CCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAAAAGTCCAAAGGATCTCAACCTTC 1084
Db 947 CACCTTGACATCTTCAACCCCTCAAGCTGGTAGCATCAAAACCGTACCGAGCTCGACTTC 1006
QY 1085 TAATACTTAGGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGGAATGCATTGT 1144
Db 1007 CAGCCCTCTCGTGGCTCAAACTCAGTGCACGAGTTTGATCCTCCGCAAGATGCTATGT 1066
QY 1145 TTCTCGCTCACTAAACACCAACGACACACAGCATCATATATCATATGATTCAGGGGACGGGCTC 1204
Db 1067 TCGTGCCACACTACAACTGAACGCAACAGCATATATACGCATTGAATGGAGCGGCAT 1126
QY 1205 ACGTCAAGTCGTGGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAGGGTC 1264
Db 1127 TGGTCAAGTGGTGAATTTGCAATGGTGCAGAGAGTGTTCGATGAGAGCTGCAAGAGGGAC 1186
QY 1265 ACCTGCTTGTGTGGCCACAGAACTTCGCGTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1324
Db 1187 AGGTGTTAATTGTGGCCACAAACTTTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1246
QY 1325 AATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAAACT 1384
Db 1247 AGTATGTTTCATTCAAGACCAATGATAGACCTCGATCGCAACCTTGAGGTGCAAACT 1306
QY 1385 CCCTCATAGATAACCTGCCGAGAGGTGGTTCGAAATTCATATATGCTCCAAAGGGAGC 1444
Db 1307 CATTGTTGAACGCAATTCGCGAGGAAAGTATTCAGCAAACTTTTAACTTAAGGAGGCAGC 1366
QY 1445 AGCAAGGAGCTTAAGCAACAAACCCCTTCAAGTTCCTTCGTTCCACCGTCTCAGCAGT 1504
Db 1367 AGCCAGGAGGAGTCAAGAAACAAACCCCTTTCAGCTTCCTGGTTCACCTAAGGAGTCTC 1426
QY 1505 CTCGAGGGCTGTGGCTTA 1523
Db 1427 AGAGGAGGTTGTGGCTTA 1445

RESULT 10
US-10-757-074-13
; Sequence 13, Application US/10757074
; Publication No. US20040139502A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10757, 074
; FILING DATE: 14-Jan-2004
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09758, 652
; FILING DATE: 11-Jan-2001
; APPLICATION NUMBER: 60/019, 940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071--A

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-992-5481
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1446 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-10-757-074-13

Query Match 34.0%; Score 517.4; DB 19; Length 1446;
Best Local Similarity 62.8%; Pred. No. 1.7e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

QY 5 AGCAACCGAGAGAGACGGTGCAGATTCCAGCGCTCAATCGCGAGAGACCTGACAATC 64
Db 71 AGCAGCCACAGCAAAACGAGTGCAGATCCAAACGCTCAATCGCTTAAACCGGATAAC 130
QY 65 GCATTGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTCAAT 124
Db 131 GTATAGATCAGAGGTGCTTCATTGAGACATGGAACCCCTAACAAAGCCATTCCAGT 190
QY 125 GCGCGGGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCTTCT 184
Db 191 GTGCGGGTGTGCCCTCTCTCGCTCACCCCTCAACCGCAACGCCCTTGCAGACCTTCT 250
QY 185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACCTTTGGTGTGATAT 244
Db 251 ACACCAACGCTCCCGAGGAGATCTATCCAAACAGGTAGTGTATTTTGGCATGATAT 310
QY 245 TCCCTGGTGTCTAGACACTATGAAGAGCTTCACACACAAGGTGCTCGATCTCAGTCCC 304
Db 311 TCCCGGTTGCTTAGCACATTGAGAGCCT----- 342
QY 305 AAAGACCACCAAGAGCTCTCAAGGAGAGAGACCAAGGCCAACAGCAACGAGATAGTCAAC 364
Db 343 -----CAACAAAAAGGACAAAGCAGCAGGCGCCCAAGACCGTCAAC 382
QY 365 AGAAGGTGACCGTTTCGATGAGGGTGATCTCATTTGCGAGTTCACCGCGTGTGCTTTCT 424
Db 383 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTGCGAGTGCACCGGTTTTGCATAT 442
QY 425 GGTCTTACAAACGACACGACACTGATGTTGTTGCTGTTTCTCTACTGACACCAACA 484
Db 443 GGATGTACAAATGAAGACACTCTCTGTTGTTGCGGTTTCTCTATTGACACCAACAGCT 502
QY 485 ACGACAACCGAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGGAACACGGAGCAAG 544
Db 503 TCCAGAACCGCTCGACAGATGCTTAGGAGATTCTATCTTCTGGGAACCAAGAGCAAG 562
QY 545 AGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGAGCAAGAAAGCTTTACCATATAGCC 604
Db 563 AGTTTCTACAGTATCAGCCACAGAAAGCAGCAAG----- 595
QY 605 CATACAGCCGCAAGTCAAGCTAGCAAGAGAGGCTGAATTTAGCCCTCGAGGACAGC 664
Db 596 -----GAGGTACTC 604
QY 665 ACAGCCGAGAGACGAGCAGGACGAAGAAGAAAGAAAGAGGTGGAAACATCTTTCAGCG 724
Db 605 AAAGCCAGAAAGGAAGCGTCAGCAAGAAGAAGAAAGAAAGAGGAGCAGCATATTGAGTG 664
QY 725 GCTTCAGCGCGAGTTCTTGGAAACAAAGCTTCAGAGTTGACGACAGACAGATAGTGCATA 784
Db 665 GCTTCGCGCGGAAATTTTGGAAACATGCGTTT-----GTGTCGACAGGAGGAGGAG 721
QY 785 ACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGGAGGAGCCATTGTGACAGTGAAGGAGGCC 844
Db 722 AGCTCAAGGTGAGAACGAAAGGAGAGAGAGGTTGCCATTGTGACAGTGAAGGAGGTC 781

Qy	485	ACGACAAACAGCTTGATCAGTTCCTCCAGCAGATTCATATTTGCTGGGAAACACGAGCAAG	544
Db	503	TCCAGAACCCAGCTCGACAGATGCTAGGAGATTCATCTCTGCTGGGAAACCAAGAGCAAG	562
Qy	545	AGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTTACCATATAGCC	604
Db	563	AGTTTCTACAGTATCAGCCACAGAGCAGCAAG-----	595
Qy	605	CATACAGCCCGCAAGTTCAGCTTAGCAAGAGAGCGTGAAATTGATGCCCTCGAGGACAGC	664
Db	596	-----GAGGTACTC	604
Qy	665	ACAGCCGACAGAAACGACGAGCAAGAAAGAAAGAAACGAAGGTGGAAACATCTTTACGCG	724
Db	605	AAAGCCAGAAAGAAAGCGTCAGCNAAGAAAGAAAGAAACGAAGGAGGACGATATTTGATGTG	664
Qy	725	GCTTCAAGCCCGAGTTCCTGAAACAAAGCTTTCCAGTTTGACGACAGACAGATAGTGCAAA	784
Db	665	GCTTCGCCCGCGAATCTTCTGAAACATGCTTC--GTCGTGACACAGGCAGATAGTAGAA	721
Qy	785	ACCTAAGAGGCGACACGAGAGTGAAGAGAGGAGCCATTTGTCACAGTGAAGGGAGGCC	844
Db	722	AGCTACAAGGTGTAGAACGAGAGGAGAAAGGGTGCCATTTGTGACAGTGAAGAGGAGTCT	781
Qy	845	TCAGAAATCTTGAGCCACAGATAGAAGAGACGCTGCCAGCAAGAGAGAGGAATACCATCAAG	904
Db	782	TCAGCGTGATTAAGCCCAACCCACGGAAGCAGCAACAAAGACCCGAGGAGAGGAGCAAGC	841
Qy	905	ATGAATATGAATACGATGAAGAGGATAGAAGCGCTGGCAGGGGAAGCAGAGGCGAGGGGA	964
Db	842	CAGATTTGTGACGAGAAACACAAACAT-----TGCCAAAGCCAAAGACGAA	886
Qy	965	ATGTTATTGAAGAGACGATCTGCACCGCAGCTGTAAAGAAACATTTGGTAGAACACAGAT	1024
Db	887	ATGGCATTGACAGAGACCATTTTGCAATATGAGACTTCGCCACAACTTTGGCCAGACTTCAT	946
Qy	1025	CCCCTGACATCTACAACCCCTCAAGCTTGGTTCACTCAAAACCTGCCAAACGATCTCAACCTTC	1084
Db	947	CACCTGACATCTTCAACCCCTCAAGCTGGTAGCATCACAAACCGCTACCAGCCTCGACTTCC	1006
Qy	1085	TAATACTTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGGAATGCATTGT	1144
Db	1007	CAGCCCTCTCGTGGCTCAAACTCAGTGGCCAGTTTGGATCACTCCGCAAGAAATGCTATGT	1066
Qy	1145	TTGTGCGTCTACTACAAACCAACGACACACAGCATCATATATCGATTGAGGGGACGGGCTC	1204
Db	1067	TCGTGCCACACTACAACCTGAACCGCAACAGCATATATACGATTAATGACGGGCAT	1126
Qy	1205	ACGTGCAAGTCGTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC	1264
Db	1127	TGTTACAAGTGGTGAATTTGCAATGGTGAGAGAGTGTTTGATGGAGAGCTGCAAGAGGGAC	1186
Qy	1265	ACGTGCTTGTGGTGCCACAGAACTTCGCGCTGGGAAAGTCCACAGCGGAGAACTTCG	1324
Db	1187	AGTGTTAATTTGTGCCACAAAACCTTTGGGTGGCTGCAAGATCACAGAGGCAAACTTCG	1246
Qy	1325	AATACGTGGCATTCAAGACAGACTCAAGGCCCAGCATAGCCAACTTCGCCGGTGAATAACT	1384
Db	1247	AGTATGTTTCATTCAAGACCAATGATAGACCTCGACCGCAACCTTTGCAAGGTGCAAACT	1306
Qy	1385	CCGTATAGATAAACCCTCCGGGAGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGGGAGC	1444
Db	1307	CATTGTTGAACGCATTCGGGAGAGAGTGATTCAGCAAACTTTTAAACCTTAAGGAGGAGC	1366
Qy	1445	AGGCAAGCGAGCTTTAAGAAACAAACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGT	1504
Db	1367	AGGCCAGCGAGTCAAGAACAAACACCCCTTCAGTTCTCTGGTTCCACCTAAGGAGTCTC	1426
Qy	1505	CTCCGAGGCGCTGGGCTTA	1523
Db	1427	AGAGGAGAGTTGTGGCTTA	1445

RESULT 13

```

US-10-425-114-7280
; Sequence 7280, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Jiongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 7280
; LENGTH: 1673
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700645818_FLI
US-10-425-114-7280

```



```
QY 665 ACAGCCGACGAGACGAGCAGGACCAAGAGAGAAACGAAAGGTGGAAACATCTTCAGCG 724
Db 627 AAAGCCAGAAAGGAAAGCGTCAGCAAGAGAGAAAGAAAGGAGGAGCAGCATATTGAGTG 686
QY 725 GCTTCAGCGCGGAGTTCCTGGACACAGCCCTTCAGGTTGACGACAGCAGATAGTCAAA 784
Db 687 GCTTCGCCCCGGAATCTTGGAACATGCGTTC---GTCGTTGGACAGCGCAGATAGTGAGAA 743
QY 785 ACCTAAGAGCGAGACGAGAGTGGAAGAGGAGGACCATTTGTGACAGTGGAGGAGGCC 844
Db 744 AGCTACAAGGTGAGACGAGAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 803
QY 845 TCAGAAATCTTTAGCCGACAGATAGAAAGAGACGTCGCCGACGAAAGAGGAAATACGATCAAG 904
Db 804 TCAGCGTGATAAGCCACCCACCGAGAGCAGCAACAAAGACCCGAGGAGAGGAGGAGGAGG 863
QY 905 ATGAATATGATACGATGAGAGAGATAGAGGCGTGGCAGGAGGAGGAGGAGGAGGAGGAGG 964
Db 864 CAGATTGTGACGAGAAAGACAAACAT-----TGCCAAAGCCAAAGCAGAA 908
QY 965 ATGGTATTGAAGAGACGATCTGCA CGCAAGTGTCTAAAGAAACATTTGGTAGAAACAGAT 1024
Db 909 ATGCGATTGACGAGACCAATTTGACATGAGACTTCGCCACAAATTTGCCAGACTTCAT 968
QY 1025 CCCTGACATCTTAAACCCCTCAAGCTGGTTTCACTCAAAAACGTCGCAACGATCTCAACCTTC 1084
Db 969 CACTGACATCTTCAACCCCTCAAGCTGGTAGCATCAACCCGTACCAGCCTCGACTTC 1028
QY 1085 TAATACTTAGTGGCTTGGACCTTAGTGTGAATATGGAATCTCTACAGGAATGCAATGT 1144
Db 1029 CAGCCCTCTCGTGGCTCAAACTCAGTGCAGCCAGTTTGGATCACTCCGCAAGAATGCTATGT 1088
QY 1145 TTGTGCTACTACAAACACCAAGCAGCAGACATCATATATCGATTGAGGGAGCGGCTC 1204
Db 1089 TCGTGCCACACTAACCTGAACGCAACAGCATATATACGCAATTTGAATGAGCGGCAT 1148
QY 1205 ACCTGCAAGTGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC 1264
Db 1149 TGGTACAAGTGTGTAATTGCAATGGTGAGAGAGTGTGATGAGAGAGCTGCAAGAGGGAC 1208
QY 1265 ACCTGCTTGGTGGCCACAGAACTTCGCGTCCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1324
Db 1209 AGTGTGTAATTTGTGCCACAAACTTTGCGGTGGCTGCAAGATCACAGAGCGCAACTTCG 1268
QY 1325 AATACTGGCATTAAGACAGACTCAAGGCCAGCAGTAGCAACCTCGCGGTGAAAACT 1384
Db 1269 AGTATGTTTCATTCAAGACCAATGATAGACCTTCGATCGCAACCTTCGAGGTGCAAACT 1328
QY 1385 CCGTCATAGATAACCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCCTCCAAAGGGAGC 1444
Db 1329 CATTGTTGAACGCTTGGCGGAGGAGTGATTTCAGCAAACTTTAACCTTAAGGAGGCAGC 1388
QY 1445 AGCAAGGAGCTTAAGAACCAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGT 1504
Db 1389 AGCCGAGGAGGTCAAGAACCAACACCCCTTTCAGCTTCCTGGTTCCACCTTAAGGAGTCTC 1448
QY 1505 CTCGAGGGCTGTGGCTTA 1523
Db 1449 AGAGGAGAGTTGTGGCTTA 1467
```

RESULT 14
US-10-425-114-8531
; Sequence 8531, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8531
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700757865_FLI
US-10-425-114-8531

Query Match 34.0%; Score 517.4; DB 18; Length 1689;
Best Local Similarity 62.8%; Pred. No. 1.8e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;

```
QY 5 AGCAACCGGAGGAGAAACGCTGCCAGTTCACGCGCTCAATGGCGAGAGACCTGACAATC 64
Db 105 AGCAGCCACAGCAAAACAGAGTGCAGATCCAAACGCTCAATGCTTCAAAACCGGATAAC 164
QY 65 GCATTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTGCAAT 124
Db 165 GTATAGAGTCAGAGGTGGCTTTCATTGAGACATGAAACCCCTAACCAACAGCCATTCCAGT 224
QY 125 GCGCGGCGTCCCTCTCTCGCTTAGTCTCGCGCAACGCGCTTCGTAGGCGCTTCT 184
Db 225 GTGCGGGTGTGCGCTCTCTCGCTGCA CCTCAACGCGCAACGCGCTTCGCAACCTTCT 284
QY 185 ACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGTTGATAT 244
Db 285 ACACCAACGCTCCCGAGGAGATCTACATCCAAACAGGTAGTGGTATTTTGGCATGATAT 344
QY 245 TCCCTGTTGCTAGACACTATGAGAGCTTCACACAAAGTCTGTCATCTCAGTCCC 304
Db 345 TCCCGGGTGTCTTAGCAATTGGAAGAGCT----- 376
QY 305 AAAGACCAACCAAGACGCTCCCAAGAGAGAACCAAAAGCCCAACAGCAACGAGATAGTCACC 364
Db 377 -----CAACAAAGGACAAAGCAGCAGGCCCCCAAGCCGTCACC 416
QY 365 AGAAGTGACCGTTCGATGAGGTGATCTCATTTGAGTTCCTCCACCGGTGTGCTTCT 424
Db 417 AGAAGATCTATCACTTCAGAGAGGGTGATTTGATTGCAAGTGCACACGCGTTTGCATACT 476
QY 425 GCTCTTCAACACGACACGACACTGTGTTGTTGCTTCTTCTTACTGACACCAACA 484
Db 477 GGATGTACAACAAATGAAGACACTCTCTGTTGTTGCGCTTCTCTTATTGACACCAACAGCT 536
QY 485 AGCAACACCGCTTGTATCAGTTCCCGAGGAGATTCAATTTGGCTGGGAACACGAGCAAG 544
Db 537 TCAGAACGAGCTTCACGACGATGCTTAGGAGATTCTATCTTCTGGGAACCAAGAGCAAG 596
QY 545 AGTTCTTAAGGTACAGCAACCAAGCAGACAAAGCAGACGAAAGAGCTTACCATATAGCC 604
Db 597 AGTTTCTACAGTATCAGCCACAGAAAGCAGCAAG----- 629
QY 605 CATACAGCCGCAAAAGTACGCTCAGCAAGAGAGGGTGAAATTTAGCCCTCAGAGACAGC 664
Db 630 -----GAGGTATCTC 638
QY 665 ACAGCGCAGAGAAACGAGCAGGACAGAAAGAAAGAAACGAAAGGTGGAACATCTTTCAGCG 724
Db 639 AAAGCAGAAAGAAAGCGTCAGCAAGAAAGAAAGAAACGAAAGGAGGAGCAGCATATTGAGTG 698
QY 725 GCTTCAGCGCGGAGTTCCTGGAAACAGCTTCAGGTTGACGACAGACAGATAGTGCAAA 784
Db 699 GCTTCGCCCCGGAATCTTGGAAACATGCGTTC---GTGTCGACAGGCGAGATAGTGAGAA 755
QY 785 ACCTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGGAGCCATTGTCAGTGAAGGGGAGGCC 844
Db 756 AGCTCAAGGTGAGAAACGAAAGAGGAGAGAGGGGTGCCATTGTGACAGTGAAGAGGAGGTC 815
```

845 TCAGATCTTGCAGCCGACATAGAGAGAGCGTCCGACGAGAGAGGATACCATGAG 904
Db |||||
816 TCAGCGGTATAGGCCCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 875
Qy |||||
905 ATGAATATGAATACCATGAGAGAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 964
Db |||||
876 CAGATTGTGACGAG 920
Qy |||||
965 ATGTATTTGAAG 1024
Db |||||
921 ATGGCAATTTGACGAG 980
Qy |||||
1025 CCCGTGACATCTACACCCCTCAAGCTGGTTCACCTCAAACTGCGCAAGAGAGAGAG 1084
Db |||||
981 CACCTGACATCTTCAACCCCTCAAGCTGGTTCACCTCAAACTGCGCAAGAGAGAG 1040
Qy |||||
1085 TAATACCTTAGGTGGTCTGAGCTAGTCTGAGTGAATATGGAATCTCTACAGGAATGCAATTGT 1144
Db |||||
1041 CAGCCCTCTCTGGCTCAAACTCAGTGGCCAGTTTGGATCACTCCGCAAGATGCTATGT 1100
Qy |||||
1145 TTGTGCTGCTACTACACCAACCGCACAGCATCATATATCGATTGAGGGGAGCGGCTC 1204
Db |||||
1101 TCGTGCCACACTACAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 1160
Qy |||||
1205 ACCTGCAAGTCTGTGACAGCAACCGCACAGAGTGTACGAGGAGAGAGAGAGAGAG 1264
Db |||||
1161 TGGTACAAAGTGGTGAATTTGCAATGGTGGAGAGAGTGTGATGGAGAGCTGCAAGAGGGAC 1220
Qy |||||
1265 ACCTGCTGTGGTGGCCACAGAACTTTCGCGCTGCGTGGAAAGTCCGAGAGAGAGAGAG 1324
Db |||||
1221 AGGTGTTAAATGTGCGCAAAACCTTTGGGTGGCTGCGAGATCAAGAGAGAGAGAG 1280
Qy |||||
1325 AATACGTGGCAATCAAGACAGACTCAAGGGCCAGCATAGCCAACTCCGCGGTGAAAACT 1384
Db |||||
1281 AGTATGTTTCAATCAAGACCAATGATAGACCTCGATCGGCAACTTTGCAAGGTGCAAACT 1340
Qy |||||
1385 CCGTCATAGATAACTCCGAGAGAGAGTGGTTCGAAATTCATATGGCTCCAAAGGAGC 1444
Db |||||
1341 CATTTGTTGAAACGCAATTCGCGAGAGAGTGTTCAGCAAACTTTTAACTTAAGGAGGAGC 1400
Qy |||||
1445 AGCAAGGAGCTTAAGAAACAAACCCCTTCAAGTTCTTCTGTTCCACCGCTCTCAGCAGT 1504
Db |||||
1401 AGGCAAGGAGCTTAAGAAACAAACCCCTTCAAGTTCTTCTGTTCCACCGCTCTCAGCAGT 1460
Qy |||||
1505 CTCGAGGGCTGTGGCTTA 1523
Db |||||
1461 AGAGGAGAGTGTGGCTTA 1479

RESULT 15

US-10-425-114-8592
; Sequence 8592, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8592
; LENGTH: 1694
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700763253_FU1
US-10-425-114-8592

Query Match 34.0%; Score 517.4; DB 18; Length 1694;
Best Local Similarity 62.8%; Pred. No. 1.9e-154;
Matches 954; Conservative 0; Mismatches 421; Indels 144; Gaps 4;
Qy 5 AGCAACCGGAGGAGAAAGCGTGCAGTTCAGCGCCCTCAATGCGCAGAGAGAGAGAGAG 64
Db 103 AGCAGCCACAGCAAAAGAGTGCAGATCCAGCCCTCAATGCCCTCAAAACCGGATACCC 162
Qy 65 GCATTGAATCAGAGGGCGGTTCATTGAGACTTTGGAAACCCCAACCAACAGAGAGTTCGAAT 124
Db 163 GTATAGAGTCAAGAGGTGGCTTCATTGAGACATGGAACCCCTAACAACAAGAGAGTTCAGT 222
Qy 125 GCGCCGGCTGCGCCCTCTCTCGCTTAGTCTCCGCGCAGAGAGAGAGAGAGAGAGAGAG 184
Db 223 GTGCGCGTGTGCGCTCTCTCGCTGACCCCTCAACCGCAACGCGCTTCGAGAGAGCTTCCT 282
Qy 185 ACTCCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGAGAGAGAGAGAGAGAGAGAG 244
Db 283 ACACCAACGCTCCCGAGGAGATCTTCATCCAAAGAGTGTGATTTATTTTGGCATGATAT 342
Qy 245 TCCTGCTGTCTCTAGACACTATGAAGAGCTTCACACACAAGAGTGTGATCTCAGTCCC 304
Db 343 TCCCGGTTGTCTAGCACATTTGAAGAGCT----- 374
Qy 305 AAAGACCAACAGAGCGTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
Db 375 -----CAACAAAAG 414
Qy 365 AGAAGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCAGCGGTGCTTCTTCT 424
Db 415 AGAAGATCTATCACTTCAG 474
Qy 425 GGTCTACAAACGACAGACACTGATGTTGCTGCTGCTTCTTCTTACTGACACCAACAACA 484
Db 475 GGATGTACAAACATGAAGACACTCTGTTGCTGCTGCTTCTTCTTATGACACCAACAAGCT 534
Qy 485 ACACAAACGAGTGTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAAACCGAGCAAG 544
Db 535 TCCAGAAACGAGTTCGACAG 594
Qy 545 AGTTCTTACAGTACAGCAACAG 604
Db 595 AGTTCTTACAGTACAGCAACAG 627
Qy 605 CACAGAGCGCAAGAGTCCAG 664
Db 628 -----GAGTACTC 636
Qy 665 ACAGCCGAG 724
Db 637 AAAGCCAG 696
Qy 725 GCTTCAGCGCGAGAGTTCCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
Db 697 GCTTCGCGCGAGAGTTCCTGGAG 753
Qy 785 ACCTAAGAGCGAG 844
Db 754 AGTCAAGAGTGAAG 813
Qy 845 TCAGAAATCTTGAGCCAG 904
Db 814 TCAGCGTGATAAGCCCAACCCAG 873
Qy 905 ATGAATATGAATACGATGAAG 964
Db 874 CAGATTGTGACGAG 918
Qy 965 ATGTATTGAAG 1024
Db 919 ATGGCAATGACGAG 978

Qy	1025	CCCTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAAAGTGCAGGATCTCAACCTTC	1084
Db	979		1038
Qy	1085	TAATACTTAGGTGGCTTGGACCTAGTGTGAATAATGGAATCTCTACAGGAATGCATTGT	1144
Db	1039		1098
Qy	1145	TTGTGCTCACTACAACCAACGACACACAGCATCATATATCGATTGAGGGGACGGGCTC	1204
Db	1099		1158
Qy	1205	ACGTGCNAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTC	1264
Db	1159		1218
Qy	1265	ACGTGCTTGTGGTGCCACAGAACTTCGCGTCTGGAAAGTCCACAGAGCGAACTTCG	1324
Db	1219		1278
Qy	1325	AATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAAAACT	1384
Db	1279		1338
Qy	1385	CCGTATAGATAACCTGCCGGAGGAGTGTGCAAAATTCATATGGCCTCCAAAAGGGAGC	1444
Db	1339		1398
Qy	1445	AGCCAGGCAGCTTAAACAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGT	1504
Db	1399		1458
Qy	1505	CTCCGAGGGCTGTGGCTTA	1523
Db	1459		1477

Search completed: August 24, 2005, 09:47:23
Job time : 931.248 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:44:21 ; Search time 86.3113 Seconds
(without alignments)
13658.075 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 CGSCACCAACCGAGGAGAA.....ctccgagggtgtggttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10728323/runat_23082005_124355_29200/app_query.fasta_1.4757
-DB=A_Geneseq_16Dec04 -OFMT=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN_1_1_224 @runat_23082005_124355_29200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2673	99.3	510	6	ABU52484
2	2673	99.3	510	7	ADG27544
3	2667	99.1	510	4	AAU04708
4	2665	99.0	507	8	ADM12139
5	2663	99.0	510	2	AAI15246
6	2652	98.6	512	2	AAI40912
7	2638	98.0	526	3	AAI33601
8	2638	98.0	526	4	AAU04711
9	2638	98.0	526	4	AAU05036
10	2615	97.2	507	8	ADO38357

11	1502	55.8	481	5	ABG71266	Abg71266 Glycine m
12	1502	55.8	481	7	ADH89253	Adh89253 G. max gl
13	1502	55.8	481	7	ADL90187	Adl90187 Soybean g
14	1502	55.8	481	8	ADG43988	Adg43988 G. max gl
15	1499.5	55.7	485	5	ABG71265	Abg71265 Glycine m
16	1499.5	55.7	485	7	ADH89247	Adh89247 G. max gl
17	1499.5	55.7	485	7	ADL90186	Adl90186 Soybean g
18	1499.5	55.7	485	8	ADG43982	Adg43982 G. max gl
19	1496	55.6	495	3	AAH80994	Aah80994 Soybean g
20	1496	55.6	495	5	ABG71264	Abg71264 Glycine m
21	1496	55.6	495	7	ADH89245	Adh89245 G. max gl
22	1496	55.6	495	7	ADL90168	Adl90168 Soybean g
23	1496	55.6	495	8	ADG43980	Adg43980 G. max gl
24	1496	55.6	511	7	ADL90190	Adl90190 Soybean g
25	1492.5	55.5	485	6	ABU52502	Abu52502 Soybean g
26	1492.5	55.5	485	7	ADG27563	Adg27563 Soybean G
27	1488	55.3	484	2	AAH40949	Aah40949 Soybean g
28	1466	54.5	495	4	AAE10365	Aae10365 Soybean g
29	1034.5	38.4	457	8	ADO43062	Ado43062 Cashew nu
30	995.5	37.0	484	1	AAH71081	Aah71081 Sequence
31	986	36.6	517	5	ABG71267	Abg71267 Glycine m
32	986	36.6	517	7	ADL90189	Adl90189 Soybean g
33	973.5	36.2	562	7	ADH89249	Adh89249 G. max gl
34	973.5	36.2	562	7	ADL90188	Adl90188 Soybean g
35	973.5	36.2	562	8	ADG43984	Adg43984 G. max gl
36	946.5	35.2	562	5	ABG71268	Abg71268 Glycine m
37	941	35.0	561	1	AAH61363	Aah61363 Soybean g
38	926.5	34.4	516	8	ADH89251	Adh89251 G. max gl
39	926.5	34.4	516	8	ADG43986	Adg43986 G. max gl
40	914	34.0	185	2	AAH40951	Aah40951 Ara h 3 a
41	901.5	33.5	516	1	AAH61362	Aah61362 Soybean g
42	891	33.1	291	8	ADG60333	Adg60333 Cholester
43	845.5	31.4	499	7	ADC08275	Adc08275 Rice prot
44	841	31.3	499	7	ADC08135	Adc08135 Rice prot
45	841	31.3	499	7	ADC08279	Adc08279 Rice prot

ALIGNMENTS

RESULT 1

ABU52484

ID ABU52484 standard; protein; 510 AA.

AC ABU52484;

DT 10-MAR-2003 (first entry)

XX Peanut Ara h3 protein sequence.

XX Peanut Ara h3 protein sequence.

XX Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

XX anaphylactic food allergen; anti-allergenic; vaccine; wound healing.

XX Arachis hypogaea.

XX WO200274250-A2.

XX 26-SEP-2002.

XX 18-MAR-2002; 2002WO-US009108.

XX 16-MAR-2001; 2001US-0276822P.

XX 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;

XX Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

XX RabbJohn PA, Shain DS, Stanley JS;

XX WPI; 2003-018765/01.

XX N-PSDB; ABX70612.

XX New modified anaphylactic food allergen, useful for preventing or

treating allergic reactions associated with e.g. anaphylactic allergens.

Claim 27; Fig 68B; 300pp; English.

The invention relates to a modified anaphylactic food allergen has an amino acid sequence that is substantially identical to that of natural anaphylactic food allergen, except for a cysteine residue that has been modified so that it cannot participate in the disulphide bond. The modification may also comprise mutation of the IgE binding sites to reduce allergenicity. Also included are: (1) a method of making a modified anaphylactic food allergen; (2) a nucleotide molecule encoding for or causing a site specific mutation in the modified anaphylactic food allergen; (3) a transgenic plant or animal expressing the modified anaphylactic food allergen; (4) a method of treating an individual by reducing the clinical response to a natural anaphylactic food allergen; and an isolated fragment of peanut allergen Ara h 1. The modified anaphylactic food allergen is useful for preventing or treating allergic reactions associated with any natural allergen such as food, insect, rubber or preferably anaphylactic allergens. It is also useful for treating wounds in mammals such as bovine, canine, feline, caprine, ovine, porcine, murine or equine species. The present sequence is a peanut allergen (e.g. Ara hl, h2 or h3)

Sequence 510 AA;

Alignment Scores:
Pred. No.: 3.77e-236 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: Gaps: 0

US-10-728-323-3 (1-1524) x ABUS2484 (1-510)

QY	1	CGCAGCAACCGAGAGAACGGCTGCGAGTTCAGCGCCTCAATCGCGAGAGACCTGAC	60
DB	4	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	23
QY	61	AATCGCATTGAATCAGAGGCGGTACATTGACACTTGGAACTCCCAACCAACGAGGATTC	120
DB	24	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	43
QY	121	GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGCAACGCCCTTCGTAGGCCT	180
DB	44	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	63
QY	181	TTCTACTCNAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTG	240
DB	64	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	83
QY	241	ATATTCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAAGGTGCTCGATCTCAG	300
DB	84	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	103
QY	301	TCCAAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCCAACAGCAACGAGATAGT	360
DB	104	SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	123
QY	361	CACCAGAGGTGACCGTTTCGATAGGGTGATCTCATTCGAGTCCCAACCGGTGTGCT	420
DB	124	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
QY	421	TTCTGGCTCTACACCAACCAACGACACTGATGTGTGTGCTGTCTCTACTGACACCAAC	480
DB	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
QY	481	AACAACGACCAACGAGTTGATCAGTTCCCGAGGAGATTCAATTTGGCTGGGAACACGGAG	540
DB	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
QY	541	CAAGAGTTCTTAGGTAACCGCAACAAAGCAGACAAAGCAGAGCAAGAGCTTACCATAT	600

DB	184	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	203
QY	601	AGCCCATACAGCCCGCAAGAGTCTAGACAAGAGAGCGTGAATTTTAGCCCTCGAGGA	660
DB	204	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	223
QY	661	CAGCAGCCGCGAGAGAACGACGACGACGAAGAGAAAGAAACGAAGGTGGAAACATCTTC	720
DB	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	243
QY	721	ACGGCTTCACCGCGAGTCTCTGGAAACAGCCCTTCCAGGTTCACGACAGATAGTG	780
DB	244	SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	263
QY	781	CAAAACCTTAAGAGGCGAGACCCAGAGTGAAGAGAGGAGGCCATTGTGACAGTGGGGA	840
DB	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	283
QY	841	GGCCTCAGAAATCTTGAGCCCGACATAGAAAGACAGCGTCCCGACGAAGAGGAATACGAT	900
DB	284	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp	303
QY	901	GAAGATGAATATGAATCAAGATCAAGAGGATGAAGCGCTGGCAGGGAAGCAGGCAGG	960
DB	304	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	323
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAAGAACATTGCTAGAAAC	1020
DB	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	343
QY	1021	AGATCCCTTGACATCTACAACCTCAAGCTGTTCACTCAAAACTGCCAACGATCTCAAC	1080
DB	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	363
QY	1081	CTTCTAATACTTAGTGGCTTGCAGCTAGTCTGATATATGGAATCTCTACAGGAATGCA	1140
DB	364	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
QY	1141	TTGTTTGTGCTCACTACAACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
DB	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
QY	1201	GCTACGTGCAAGTCTGGACCAACGCAACAGAGTGTACGACGAGAGCTTCAAGAG	1260
DB	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	423
QY	1261	GCTACGTGCTGCTGGTCCACAGAACTTCGCCCTCGCTGGAAAGTCCGACGAGCGAGAC	1320
DB	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
QY	1321	TTTCAATACGTGGCATTCAAGACAGACTCAAGGCCCGACATAGCCAACCTCGCCGTGAA	1380
DB	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
QY	1381	AACTCCGTATAGATAACTCGCGAGGAGTGTGCAAAATTCATATATGCTCCAAAGG	1440
DB	464	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	483
QY	1441	GAGCAGGCGAGCAGCTTAAGNACAACAACCCCTCAAGTTCCTGTTCCACCGCTCAG	1500
DB	484	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln	503
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	504	GlnSerProArgAlaValAla	510
RESULT 2			
ID	ADG27544	standard; protein; 510 AA.	
XX	ADG27544		
AC	ADG27544;		
XX	26-FEB-2004	(first entry)	

XX Peanut allergen Ara h3.
 XX
 XX
 KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycine A2B1a; Jug n1;
 KW antiallergic; vulnerary; anaphylactic food allergen; Igs; allergy; wound.
 XX
 XX Arachis hypogaea.
 XX
 XX US2003202980-A1.
 XX
 XX 30-OCT-2003.
 XX
 XX 18-MAR-2002; 2002US-00100303.
 XX
 XX 29-DEC-1995; 95US-0009455P.
 XX 23-SEP-1996; 96US-00717933.
 XX 31-JAN-1998; 98US-0073283P.
 XX 13-FEB-1998; 98US-0074590P.
 XX 13-FEB-1998; 98US-0074624P.
 XX 13-FEB-1998; 98US-0074633P.
 XX 29-JUN-1998; 98US-00106872.
 XX 27-AUG-1998; 98US-00141220.
 XX 13-NOV-1998; 98US-00191593.
 XX 29-JAN-1999; 99US-00240557.
 XX 29-JAN-1999; 99US-00241101.
 XX 11-FEB-1999; 99US-00248673.
 XX 11-FEB-1999; 99US-00248674.
 XX 02-MAR-1999; 99US-0122450P.
 XX 02-MAR-1999; 99US-0122452P.
 XX 02-MAR-1999; 99US-0122560P.
 XX 02-MAR-1999; 99US-0122565P.
 XX 11-MAR-1999; 99US-00267719.
 XX 28-JAN-2000; 2000US-00494096.
 XX 16-MAR-2001; 2001US-0276822P.
 XX
 XX (CAPL/) CAPLAN M J.
 XX (SOSI/) SOSIN H B.
 XX (SAMP/) SAMPSON H.
 XX (BANN/) BANNON G A.
 XX (BURK/) BURKS A W.
 XX (COCK/) COCKRELL G.
 XX (COMP/) COMPADRE C M.
 XX (CONN/) CONNAUGHTON C.
 XX (HELM/) HELM R M.
 XX (KING/) KING N E.
 XX (KOPP/) KOPPER R A.
 XX (MALE/) MALEKI S J.
 XX (RABJ/) RABJOHN P A.
 XX (SHIN/) SHIN D S.
 XX (STAN/) STANLEY J S.
 XX
 XX Caplan MJ, Sosin HB, Sampson H, Bannon GA, Burks AW, Cockrell G;
 XX Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 XX Rabbjohn PA, Shin DS, Stanley JS;
 XX
 XX WPI; 2003-875632/81.
 XX N-PSDE; ADG27543.
 XX
 XX New modified anaphylactic food allergen comprising a cysteine residue
 XX which has been modified so that it cannot participate in the disulfide
 XX bond, useful for treating allergic reactions or wounds.
 XX
 XX Example 16; SEQ ID NO 89; 194pp; English.
 XX
 XX The invention relates to a modified anaphylactic food allergen whose
 XX amino acid sequence is substantially identical to that of a natural
 XX anaphylactic food allergen. The natural anaphylactic food allergen
 XX includes at least one cysteine residue that participates in a disulphide
 XX bond when the natural anaphylactic food allergen is in its native
 XX conformation, except that the cysteine residue has been modified so that
 XX it cannot participate in the disulphide bond. Also included are a method
 XX of making a modified anaphylactic food allergen, a nucleotide molecule

CC encoding a modified anaphylactic food allergen defined above, a
 CC nucleotide molecule for causing a site specific mutation in a gene
 CC encoding a natural anaphylactic food allergen, a transgenic plant or
 CC animal expressing a modified anaphylactic food allergen defined above, a
 CC method of treating an individual by reducing the clinical response to a
 CC natural anaphylactic food allergen by administering a modified
 CC anaphylactic food allergen and an isolated fragment of peanut allergen
 CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
 CC ADG27465. About 10-17% of the amino acids have been modified in at least
 CC one IGE epitope or all the IGE epitopes recognised when the natural
 CC anaphylactic food allergen is contacted with serum IGE from individual(s)
 CC allergic to the natural anaphylactic food allergen. The invention
 CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
 CC cDNAs), Soybean Glycinin A2B1a and IGE-binding epitopes of the English
 CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
 CC used for treating allergic reactions or wounds. The present sequence
 CC represents a Peanut allergen of the invention (or its fragment).
 XX
 XX SQ Sequence 510 AA;

Alignment Scores:

Pred. No.: 3,77e-236 Length: 510
 Score: 2673.00 Matches: 507
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.33% Indels: 0
 DB: Gaps: 0

US-10-728-323-3 (1-1524) x ADG27544 (1-510)

QY 1 CGGACGACACCGAGGAGAACGGTCCAGTTCAGCGGCTCAATGGCGAGACCTGAC 60
 |||||
 DB 4 ArgGlnGlnProGluGluAenAlaCyGlnPheGlnArgLeuAenAlaGlnArgProAsp 23
 |||||
 QY 61 AATCGCATTGAATCAGAGGCGGTACATTGACACTTGGAACTCCCAACCCCAACACAGGAGTTC 120
 |||||
 DB 24 AsnArgIleGluSerGluGlyIleGluThrTriaAsnProAsnGlnGluPhe 43
 |||||
 QY 121 GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCTCGCGCAACGCCCTTCGTAGGCCT 180
 |||||
 DB 44 GluCyAlaGlyValAlaLeuSerArgLeuValLeuArgArgAenAlaLeuArgArgPro 63
 |||||
 QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAGAGGAGGATCTTTGGTTG 240
 |||||
 DB 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
 |||||
 QY 241 ATATTCCTGTTGTCTTAGACACTATGAAGAGCCTCACACAAAGTCGTCTCATCTCAG 300
 |||||
 DB 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
 |||||
 QY 301 TCCCAAGACCAACCAAGACGTCTTCAAGGAGAGAGACCAAGCCAAAGCCACGACGATAGT 360
 |||||
 DB 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
 |||||
 QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCTATTGCAGTTCACCGGCTGTGCT 420
 |||||
 DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 |||||
 QY 421 TTCTGCTCTTACAAACGACCAACGACACTGATGTTGTTGTTGTTCTTCTTACTGACCAAC 480
 |||||
 DB 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
 |||||
 QY 481 AACACGACACCAACGAGCTTGATCAGTTTCCCAGGAGATTCAATTTGGCTGGGAACACGGAG 540
 |||||
 DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
 |||||
 QY 541 CAAGATCTTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAGAGACCTTACCNAT 600
 |||||
 DB 184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
 |||||
 QY 601 AGCCCATACAGCCCGCAAAAGTACGCTTACGAAGAAGAGCGTGAATTTAGCCCTCCGAGGA 660
 |||||
 DB 204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
 |||||

Qy	661	CAGCACAGCCCGCAGAGAACCGAGCAGGCAAGAGAAACGAAAGTGTGAAACATCTTC	720
Db	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	243
Qy	721	AGCGGCTTCACGCCGGAGTTCCTGGAAACAAGCCCTCCAGGTTGACGACACAGATAGTG	780
Db	244	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspAArgGlnIleVal	263
Qy	781	CAAAACCTAAGAGCGGAGACCGAGAGTCAAGACAGGAGCCATTGTGACAGTGGGGGA	840
Db	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	283
Qy	841	GGCCTCAGAATCTTGAGCCCGAGATAGAAAGAGACGTCCTCCGACGAAGAGGAAATACGAT	900
Db	284	GlyLeuArgIleLeuSerProAspArgGlyArgAlaAspGluGluGluTyArgAsp	303
Qy	901	GAAGATGAATATGAATACGATGAAGAGATAGAGGCGTGCAGGGGAAACGACGAGGAGG	960
Db	304	GluAspGluTyGluTyArgGluGluAspArgArgGlyArgGlySerArgGlyArg	323
Qy	961	GGGAATCGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAACAATTTGGTAGAAAC	1020
Db	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleGlyAsnIleGlyArgAsn	343
Qy	1021	AGATCCCTCGACATCTCAACCCCTCAAGCTGGTTCACTCAAAACCTGCCAACGATCTCAAC	1080
Db	344	ArgSerProAspIleTyAsnProGlnAlaGlySerLeuIleThrAlaAsnAspLeuAsn	363
Qy	1081	CTTCTTAATACTTAGGTGGCTTGACCTAGTGTCTGAATATGAAATCTCTACAGAATGCA	1140
Db	364	LeuLeuIleLeuArgTyrLeuGlyProSerAlaGluTyGlyAsnLeuTyArgAsnAla	383
Qy	1141	TTGCTTCGCTCTACTACAACCAACGACACGACAGCATCATATATCGATTGAGGGGACGG	1200
Db	384	LeupheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyArgLeuArgGlyArg	403
Qy	1201	GCTCACGTCGCAAGTCGTGGCAGCAACCGCAACAGAGCTGTACGACGAGGAGCTTCAAGAG	1260
Db	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyAspGluGluLeuGlnGlu	423
Qy	1261	GGTCACGTCGTTGTGGTGCACAGAACTTCGCGCTCGCTGGAAAGTCCGACGAGGAGAAC	1320
Db	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyIleSerGlnSerGluAsn	443
Qy	1321	TTCCGAATACGTGGCATTCAAGACAGACTCAAGGCCACAGCATAGCAACCTCGCGGTGAA	1380
Db	444	PheGluTyValAlaPheIleTyThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
Qy	1381	AACTCCGCTCATAGATAACCTCCGCGAGGAGTGGTTGCCAAATTATATGGCTCCCAAGG	1440
Db	464	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyGlyLeuGlnArg	483
Qy	1441	GAGCAGCAAGCGAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG	1500
Db	484	GluGlnAlaArgGlnLeuIleAsnAsnAsnProPheIleValProProSerGln	503
Qy	1501	CAGTCTCCGAGGCTGTGGCT	1521
Db	504	GlnSerProArgAlaValAla	510
RESULT 3			
AAU04708			
ID	AAU04708 standard; protein; 510 AA.		
XX			
AC	AAU04708;		
XX			
DT	23-OCT-2001 (first entry)		
XX			
DE	Anaphylactic antigen Ara h 3.		
XX			
KW	Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;		
KW	allergy; mast cell; basophil; mouse.		

Db	64	SerSerAenAlaProGlnGluIlePhe GlnGlnGlyArgGlyTyrPheGlyLeu	83	Db	424	GlyHisValLeuValValProGlnAenPheAlaValAlaGlyLysSerGlnSerGluAen	443
Qy	241	ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG	300	Qy	1321	TTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTCGCGGTGAA	1380
Db	84	IlePheProGlyCySerProArgHisTyrGluGluProHisThrGlnGlyArgSerGln	103	Db	444	PheGluTy-ValAlaPheLysThrAspSerArgProSerIleAlaAenLeuAlaGlyGlu	463
Qy	301	TCCCAAGACACCAAGACGCTCCCAAGGAGAGACCAAGGCAACAGCAACGAGATGT	360	Qy	1381	AACTCCGTCATAGATTAACCTGCCGAGGAGGTGGTTCGAAATTCATATGCGCTCCAAGG	1440
Db	104	SerGlnArgProProArgArgLeuGlnGlyLysPheGlnSerGlnGlnArgSer	123	Db	464	AenSerValIleAspAenLeuProGluGluValValAlaAenSerTyrGlyLeuGlnArg	483
Qy	361	CACCAGAAGTGACACCGTCTTCGATGAGGTGATCTCATTCAGTCTCCACCGGTGTGCT	420	Qy	1441	GAGCAGCGCAAGGCGAGCTTAAGAACCAACCCCTTCAAGTTCCTGTTCCACCGTCTCAG	1500
Db	124	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143	Db	484	GluGlnAlaArgGlnLeuLysAenAenProPheLysPhePheValProProSerGln	503
Qy	421	TTCTGCGCTTACACGACACGACACTGATGTGTGCTGTTCTCTTACTGACCAAC	480	Qy	1501	CAGTCTCCAGGCGCTGTGCT	1521
Db	144	PheTrpLeuTyrAenAspHisAspThrAspValValAlaValSerLeuThrAspThrAen	163	Db	504	GlnSerProArgAlaValAla	510
Qy	481	AACACGACAAACCAGCTTGATGATTCCTCCAGGAGATTCATTTGGCTGGGAACAGGAG	540	Db	504	GlnSerProArgAlaValAla	510
Db	164	AsnAenAspAenGlnLeuAspGlnPheProArgArgPheAenLeuAlaGlyAenThrGlu	183	Db	504	GlnSerProArgAlaValAla	510
Qy	541	CAAGAGTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600	Db	504	GlnSerProArgAlaValAla	510
Db	184	GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr	203	Db	504	GlnSerProArgAlaValAla	510
Qy	601	AGCCCATACAGCCCGCAAGCTAGCCTTAGACAAGAAGAGCGTGAATTTAGCCCTCGAGGA	660	Db	504	GlnSerProArgAlaValAla	510
Db	204	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	223	Db	504	GlnSerProArgAlaValAla	510
Qy	661	CAGCAGCCCGCAGAGAACAGCAGAGCAAGAGAGAAACGAAGGTGGAACATCTTC	720	Db	504	GlnSerProArgAlaValAla	510
Db	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGlyGlyAenIlePhe	243	Db	504	GlnSerProArgAlaValAla	510
Qy	721	AGCGGCTTACCGCGGAGTCTCGGAACAGCCCTCCAGGTTCCACGACAGACAGATGTG	780	Db	504	GlnSerProArgAlaValAla	510
Db	244	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspArgGlnIleVal	263	Db	504	GlnSerProArgAlaValAla	510
Qy	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAGGGA	840	Db	504	GlnSerProArgAlaValAla	510
Db	264	GlnAenLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	283	Db	504	GlnSerProArgAlaValAla	510
Qy	841	GGCCTCAGAACTTTGAGCCCGACATAGAAAGAGACGTGCCGACCAAGAGGAATACGAT	900	Db	504	GlnSerProArgAlaValAla	510
Db	284	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	303	Db	504	GlnSerProArgAlaValAla	510
Qy	901	GAAGATCAATATCAATACGATGAAGAGATAGAGCGTCCGAGGAGGAGCAGGAGCAGG	960	Db	504	GlnSerProArgAlaValAla	510
Db	304	GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyThr	323	Db	504	GlnSerProArgAlaValAla	510
Qy	961	GGAAATGGTATTGAAGAGACGATCTGACCCGCAAGTCTAAAGAAACATTGGTAGAAAC	1020	Db	504	GlnSerProArgAlaValAla	510
Db	324	GlyAenGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAenIleGlyArgAen	343	Db	504	GlnSerProArgAlaValAla	510
Qy	1021	AGATCCCTCAGACTCTACAAACCTCAAGCTGGTTCATCAAACTGCCCAACGATCTCAAC	1080	Db	504	GlnSerProArgAlaValAla	510
Db	344	ArgSerProAspIleTyrAenProGlnAlaGlySerLeuLysThrAlaAenAspLeuAen	363	Db	504	GlnSerProArgAlaValAla	510
Qy	1081	CTTCTAATCTAGTGGCTTGACCTAGTGGTGAATATGGAATCTCTACAGGAATGCA	1140	Db	504	GlnSerProArgAlaValAla	510
Db	364	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAenLeuTyrArgAenAla	383	Db	504	GlnSerProArgAlaValAla	510
Qy	1141	TTGTTTGTGCTCACTACACCAACGACACAGCATCATATATGATTGAGGGGACGG	1200	Db	504	GlnSerProArgAlaValAla	510
Db	384	LeuPheValAlaHisTyrAenThrAenAlaHisSerIleIleTyrArgLeuArgGlyArg	403	Db	504	GlnSerProArgAlaValAla	510
Qy	1201	GCTCAGTGAAGTCTGTGACAGCAACGCGCAACAGAGTGTACGAGGAGGCTTCAAGAG	1260	Db	504	GlnSerProArgAlaValAla	510
Db	404	AlaHisValGlnValValAspSerAsnGlyAenArgValTyrAspGluGluLeuGlnGlu	423	Db	504	GlnSerProArgAlaValAla	510
Qy	1261	GGTCAGTGTGCTGTGTCACAGAACTTCGCGTGGAAAGTCCAGAGCGAGAAC	1320	Db	504	GlnSerProArgAlaValAla	510

RESULT 4

ADM12139

ID ADM12139 standard; protein; 507 AA.

XX

AC ADM12139;

XX

DT 20-MAY-2004 (first entry)

XX

DE Arachis hypogaea 3 (Ara h3) protein.

XX

KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;

KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;

KW antiarthritic; neuroprotective; dermatological; immunosuppressive;

KW antinflammatory; antidiabetic; antithyroid; immune;

KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;

KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;

KW scleroderma; dermatomyositis; pemphigus.

XX

OS Arachis hypogaea.

XX

PN US2003235594-A1.

XX

PD 25-DEC-2003.

XX

PF 17-SEP-2002; 2002US-00245871.

XX

PR 14-SEP-1999; 99US-00396813.

PR

PR 17-JUL-2002; 2002US-00197000.

XX

XX (ANTI-) ANTIGEN EXPRESS INC.

XX

XX Humphreys R, Xu M;

PI

XX WPI; 2004-070554/07.

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC	erythematous, diabetes mellitus, myasthenia gravis, autoimmune
CC	thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC	represents a mammalian II key related protein of the invention.
XX	
SQ	Sequence 507 AA;
Alignment Scores:	
Pred. No.:	2,04e-235
Score:	2665.00
Percent Similarity:	99.80%
Best Local Similarity:	99.80%
Query Match:	99.03%
DB:	8
US-10-728-323-3 (1-1524) x ADM12139 (1-507)	
QY	1 CGGAGCAACCGAGGAGAACGGTCCAGTTCACGGCCCTCAATGCGGAGAGACCTGAC 60
DB	1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY	61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
DB	21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 40
QY	121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
DB	41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 60
QY	181 TTCTACTCCANTGCTCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTG 240
DB	61 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY	241 ATATTCCCTGGTTGCTAGACACTATGAAGAGCCTCACACACAAGTGTCTGATCTCAG 300
DB	81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 100
QY	301 TCCCAAGACCAACAGAGCTCTCCAAGGAAGAACCAAGCAACAGCAACAGAGATG 360
DB	101 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY	361 CACGAGAGTGACCGTTTCGATGAGGTGATCTCATTCGAGTTCGCCCGGTGTGCT 420
DB	121 HisGlnLysValHisArgThrAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY	421 TTCTGGCTCTACACACACACACTGATGTTGTTGCTGTTTCTTCTACTGACACCAAC 480
DB	141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY	481 AACACGACCAACGAGCTTGATCAGTTCGCCGAGAGATTCAATTTGCTGGCAACGAG 540
DB	161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY	541 CAAGAGTCTTAAGGTACGAGCAACAAAGCAGACAAAGCAGAGAGAGCTTACCATAT 600
DB	181 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 200
QY	601 AGCCCATACAGCCGCAAGTCTAGCTAGACAAGAGAGCGTCAATTTAGCCCTCGAGA 660
DB	201 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY	661 CAGCACGCGCAGAGAACGAGGAGCAAGAGAGAAAGAAACGAAGGTGGAACATCTTC 720
DB	221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 240
QY	721 AGCGGCTTCAGCGGAGTTCCTGGAAACAAGCCCTTCAGGTTTCACACAGACAGATAG 780
DB	241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY	781 CAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAGGGA 840
DB	261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
QY	841 GGCCTCAGAAATCTTGAGCCGACAGATGAAGAGAGACGTGCCCGACGAAGGAATACGAT 900

DB	281 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 300
QY	901 GAAGATGATATATGATACGATGAAGGATAGAACGGCTGGCAGGGAAGCAGAGCAGG 960
DB	301 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY	961 GCGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAAAGAACATTTGGTAGAAAC 1020
DB	321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAlaLysAsnIleGlyArgAsn 340
QY	1021 AGATCCCTTGACATCTACAAACCTCAAGCTGTTTCACTCAAAACCTGCCAACCATCTCAAC 1080
DB	341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY	1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGATGCA 1140
DB	361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380
QY	1141 TTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGGACGG 1200
DB	381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY	1201 GCTCAGTCCAAAGTCTGGACAGCAACGGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
DB	401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
QY	1261 GGTCAAGTCTGTTGGTCCACAGAACTTCGCGCTCGTGGAAAGTCCCGAGAGCGAGAAC 1320
DB	421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY	1321 TTTCGAATACGTGGCATTTCAAGACAGACTCAAGGCGCCAGCATAGCCAACTCCGCGGTGAA 1380
DB	441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY	1381 AACTCGTCTATAGATAA CCTCGCGAGGAGGTGGTTGCAAAATTCATATGGCTCCAAAGG 1440
DB	461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY	1441 GAGCAGGCAAGCAGCTTAAGAACCAACACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500
DB	481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProSerGln 500
QY	1501 CAGTCTCCGAGGCTGTGGCT 1521
DB	501 GlnSerProArgAlaValAla 507
RESULT 5	
ID	AAV15246
XX	AAV15246 standard; protein; 510 AA.
AC	AAV15246;
XX	
DT	17-OCT-2003 (revised)
DT	09-NOV-1999 (first entry)
XX	
DE	Peanut allergen, Ara h 3, amino acid sequence.
XX	
KW	allergy; immune response; transgenic; allergen; epitope;
XX	immunoglobulin E; Ig E; binding site; peanut.
OS	Arachis hypogaea.
XX	
PN	WO938978-A1.
XX	
PD	05-AUG-1999.
XX	
PF	29-JAN-1999; 99WO-US002031.
XX	
PR	31-JAN-1998; 98US-0073283P.
PR	13-FEB-1998; 98US-0074590P.
PR	13-FEB-1998; 98US-0074624P.
PR	13-FEB-1998; 98US-0074633P.

AC AAY40912;
 XX 17-OCT-2003 (revised)
 DT 06-DEC-1999 (first entry)
 XX Ara h 3 allergen sequence.
 DE Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
 KW allergic reaction; ss.
 KW Arachis hypogaea.
 OS
 XX
 XX Key Location/Qualifiers
 FH Region 187..188
 FT /note="these two amino acid residues are not indicated
 FT in the Ara h 3 sequence provided in the sequence listing
 FT (seq ID No: 6) and Fig 51 of the specification"
 FT Misc-difference 374
 FT /note="encoded by CCT"
 XX
 XX W09945961-A1.
 PN
 XX
 PD 16-SEP-1999.
 XX
 XX 12-MAR-1999; 99WO-US005494.
 PF
 XX
 PR 12-MAR-1998; 98US-0077763P.
 PR 11-MAR-1999; 99US-00077763.
 XX
 XX (UYAR-) UNIV ARKANSAS.
 PA
 XX
 XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
 PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
 PI
 XX
 DR WPI; 1999-551218/46.
 DR N-PSDB; AA222280.
 XX
 XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
 PT animal from allergic reaction.
 PT
 XX
 PS Claim 8; Fig 11; 193pp; English.
 XX
 CC The invention provides a tertiary structure for the peanut allergen Ara H
 CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
 CC epitopes. The invention also provides an isolated recombinant peanut
 CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
 CC allergen Ara h 3. Molecules of the invention are used to protect a host
 CC animal from allergic reaction, particularly using a modified allergen
 CC which is less reactive with IgE. The invention may also be used to ensure
 CC that the allergen is not introduced into genetically modified food. The
 CC present sequence represents a DNA encoding the Ara h 3 allergen. (Updated
 CC on 17-OCT-2003 to standardise OS field)
 CC
 SQ Sequence 512 AA;
 Alignment Scores:
 Pred. No.: 3,21e-234 Length: 512
 Score: 2652.00 Matches: 506
 Percent Similarity: 99.41% Conservative: 0
 Best Local Similarity: 99.41% Mismatches: 1
 Query Match: 98.55% Indels: 2
 DB: 2 Gaps: 1
 US-10-728-323-3 (1-1524) x AAY40912 (1-512)
 QY 1 CGGCAGCAACCGAGAGACCGGTGCCAGTTCACCGCCTCAATCGGCAGAGACCTGAC 60
 Db
 Db 4 ArGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
 QY 61 AATCGCATTCATACAGGCGGTACATTGAGACTTGGAAACCCCAACACCGAGAGTTC 120
 Db 24 AsnArgIleGluSerGluGlyGlyIleGluThrIleGluThrIleGluThrIleGluThr 43

QY 121 GAATGCGCGCGTCTGCTCTCTGCTTAGTCTCTCGCGCAACGCCCTTCGTAGCCCT 180
 Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
 QY 181 TTCTACTCCATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACATTTGGGTG 240
 Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
 QY 241 ATATTCCCTCGGTGCTTAGACACTATGAAGAGCCTCACACACAAGGTCGTCTCATCTCAG 300
 Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
 QY 301 TCCCAAGACCAACAGAGCTCTCCAGGAGAGACCAAGCCAAAGCAACAGACAGATAGT 360
 Db 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 123
 QY 361 CACCAAGAGGTCCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
 Db 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
 QY 421 TTCTGCTCTACAACGACCCACGACACTGATGTGTGTGCTGTTTCTTCTTACTGACACCAAC 480
 Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 163
 QY 481 AACAAACGACCAACAGCTTCATCAGTCCCGAGAGATTCATTTGGTGGGAAACAGGAG 540
 Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
 QY 541 CAA-----GAGTTCTTAAGGTACCAGCAGCAACAAAGCAGACGAGCAAGACAGCTTA 594
 Db 184 GlnGluPheGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeu 203
 QY 595 CCATATAGCCCATACAGCCCGCAAGTTCAGCTACAGCAAGAGAGCGTCAATTTAGCCCT 654
 Db 204 ProTyrSerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerPro 223
 QY 655 CGAGGACGACGACGCGCAGAGAACGAGCAGACGACGAAGAGAAAGAAACGAGGTGGAAC 714
 Db 224 ArgGlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsn 243
 QY 715 ATCTTCAGGGCTTCACGCGGAGTTCCTGGACACAGGCTTCAGGTTGACACAGACAG 774
 Db 244 IlePheSerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGln 263
 QY 775 ATAGTGCAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTACAGTG 834
 Db 264 IleValGlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrVal 283
 QY 835 AGGGAGGCGCTCAGAACTTTGAGCCCGACATAGAAAGAGACGTGCCGACGAAAGAGGAA 894
 Db 284 ArgGlyGlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGlu 303
 QY 895 TACGATGAAGATGAATATGATACCATGAGGATAGAGGCTGGAGCGGTGGCGAGGAGCAGA 954
 Db 304 TyrAspGluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArg 323
 QY 955 GGCAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTGCTGATATGGAATCTTACAG 1014
 Db 324 GlyArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleGlyAsnIleGly 343
 QY 1015 AGAAACAGATCCCTGACATCTTAAACCCCTCAAGCTGGTTCTCACTCAAACTCCCAACGAT 1074
 Db 344 ArgAsnArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAsp 363
 QY 1075 CTCACCTTCTAATACTTAGTGGCTTCGACCTAGCTGCTGATATGGAATCTTACAGG 1134
 Db 364 LeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArg 383
 QY 1135 AATGCAATTTGTTGCTCGCTCACTTACCAACCAACGACACAGCATCATATATCGATTGAGG 1194
 Db 384 AsnAlaLeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArg 403
 QY 1195 GGACGGGCTCACGTGCAAGTCTGTGGACAGCAACGGCAACAGAGTGTACGACGAGGACTT 1254

Db 404 GlyArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeu 423
 Qy 1255 CAAGAGGTCACGTGCTTGTGGTCCACAGAACTTCGCCGTGCGTGGAAAGTCCAGAGC 1314
 Db 424 GlnGluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSer 443
 Qy 1315 GAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCCACTCGCC 1374
 Db 444 GluAsnPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAla 463
 Qy 1375 GGTGAAAACTCCGTCATAGATAACCTCCGAGGAGGTGGTTCGAAATTCATATGGCCTC 1434
 Db 464 GlyLysAsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeu 483
 Qy 1435 CAAGGAGCAGCAGCAGCAGCTTAAGAAACAACCCCTTCAGTTCTTCGTTCCACCG 1494
 Db 484 GlnArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProPro 503
 Qy 1495 TCTCAGCAGTCTCCGAGGCTGGCT 1521
 Db 504 SerGlnGlnSerProArgAlaValAla 512

RESULT 7

AAB33601

ID AAB33601 standard; protein; 526 AA.

XX

AC AAB33601;

XX

XX 12-SEP-2003 (revised)

DT

22-JAN-2001 (first entry)

XX

DE Modified Ara h 3 amino acid sequence.

XX

Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;
 Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;
 food antigen; sensitising; immune response; anti-allergic.

XX

OS Arachis hypogaea.

XX

PN WO200051647-A2.

XX

PD 08-SEP-2000.

XX

PF 03-MAR-2000; 2000WO-US005655.

XX

PR 03-MAR-1999; 99US-0122960P.

XX

PR 06-DEC-1999; 99US-00455294.

XX

PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX

PI Sampson HA;

XX

DR WPI; 2000-611341/58.

XX

PT Non-human animal sensitized to an antigen, useful as an animal model for
 PT studying allergic reactions to allergens, such as those in food and in
 PT the environment.

XX

PS Example 6; Fig 17C; 124pp; English.

XX

CC The present invention describes an animal model which can be used for
 CC studying allergic reactions to allergens. The animal is sensitised to a
 CC selected antigen by administering the antigen itself or a nucleic acid
 CC encoding the antigen, where preferably the antigen is an anaphylactic
 CC antigen. The sensitised animal can then be used to screen for compounds
 CC which may help to prevent, ameliorate, or cure allergic conditions in
 CC humans. The animal model can be used for studying allergic reactions to
 CC allergens, such as those in food (peanuts, fruit, berry, nut, bean, milk,
 CC dairy products), or in the environment (weed pollen, grass pollen, tree
 CC pollen, mite, animal, animal dander, fungal, and insect antigens).
 CC AAB33478 to AAB33601 represent sequences which are used in examples from
 CC the present invention to specifically examine the peanut allergy, and the

CC

CC peanut anaphylactic antigens Ara h 1 to 3. (Updated on 12-SEP-2003 to
 CC standardise OS field)

XX

SQ Sequence 526 AA;

XX

Alignment Scores:

Pred. No.: 6,27e-233 Length: 526
 Score: 2638.00 Matches: 502
 Percent Similarity: 99.01% Conservative: 0
 Best Local Similarity: 99.01% Mismatches: 5
 Query Match: 98.03% Indels: 0
 DB: 3 Gaps: 0

US-10-728-323-3 (1-1524) x AAB33601 (1-526)

Qy 1 CGGCACCAACCGGAGGAGAACGCGTCCAGTTCACGCGCTCAATGCGCAGACCTGAC 60

Db

5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24

Qy

61 AATCGCATTCGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120

Db

25 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44

Qy

121 GAATGCGCGGCGTCCGCTCTCTCGTCTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180

Db

45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgPro 64

Qy

181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGGATCTTTGGGTG 240

Db

65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84

Qy

241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAGGTCGTGATCTCAG 300

Db

85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104

Qy

301 TCCCAAGACACCCAGAGAGTCTCCAGGAGAGAGACCAAGGCCCAACAGCAACGAGATG 360

Db

105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124

Qy

361 CACCAGAAGTGCACCGTTTTCGATGAGGTGATCTCATTCGAGTTCCTCCCGGTGTGCT 420

Db

125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144

Qy

421 TTCTGCTCTACAAACGACCACTGATGTTGTTGCTGTTGCTTCTTCTACTGACCAAC 480

Db

145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164

Qy

481 AACACGACCAACGCTTCGATCAGTTCCCGAGGAGATTCAATTTGGCTGGACACGGAG 540

Db

165 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 184

Qy

541 CAAGAGTTCCTTAAGGTACCAACCAACAGCAAAAGCAGACGAAAGAGCTTACCATAT 600

Db

185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204

Qy

601 AGCCCATACAGCCCGCAAGTCCCTAGCAAGAAAGAGCGTGAATTTAGCCCTCCAGGA 660

Db

205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224

Qy

661 CAGCAGCCGCGAGAGAACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720

Db

225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluPhe 244

Qy

721 AGCGGCTTCACCGCGAGTTCCTCGAACCAAGCCTTCAGGTTTCACGACAGACAGATAGT 780

Db

245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264

Qy

781 CAAACCTTAAGAGCGGAGACCCAGAGTGAAGAGGAGGAGCCATTGTGACGTAGGGGA 840

Db

265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGluGly 284

Qy

841 GGCCTCAGAACTTTGAGCCCGACAGACGACGACGACGACGACGACGACGACGACGAC 900

Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATGAATATACATACGATGAAGAGGATAGAGGCGTGGCAGGCGAAGCAGAGCAGG 960
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlySerArgGlyArg 324
QY 961 GGAATGGTATTGAAGAGAGCATCTCGACCGCGCAAGTGCTAAAGAACATTGGTAGAAGC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTGACATCTACACCTCAAGCTGCTCACTCAAACTGCCAACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATACTAGGTGGCTTGACCTAGTGCCTGAATATGGAATCTCTACAGGAATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
QY 1141 TTGTTTGTGCTCACTACACCAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
QY 1201 GCTACGTGCAAGTCTGACACCAACGACGAGAGTGTACACGAGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 424
QY 1261 GGTCAAGTGTGCTGTCGCACAGAACTTCGCGCTCGCTGGAAGTCCACAGAGCGAGAAC 1320
Db 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTCGAATACGTGGCATTCAGACAGACTCAAGCCAGCATGCCAACTCCGCGGTGAA 1380
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTATAGATACCTGCGGAGGAGGTGTTGCAATTCATATGCGCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuGlnArg 484
QY 1441 GAGCAGCGAAGCAGCTTAAGAACAAACACCCCTTCAAGTTCTTCCGCTCCACCGTCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProSerGln 504
QY 1501 CAGTCTCCAGGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

RESULT 8
AAU04711
XX AC AAU04711, standard; protein; 526 AA.
XX AC AAU04711;
XX DT 23-OCT-2001 (first entry)
XX DE Modified anaphylactic antigen Ara h 3.
XX KW Ara h 3; anaphylactic antigen; immunoglobulin E; IGE; immunogenic;
XX KW allergy; mast cell; basophil; mouse.
XX OS Mus sp.
XX OS Synthetic.
XX PN WO200140264-A2.
XX PD 07-JUN-2001.
XX PF 06-DEC-2000; 2000WO-US033124.
XX PR 06-DEC-1999; 99US-00455294.
XX PR 23-JUN-2000; 2000US-0213765P.
XX PR 27-SEP-2000; 2000US-0235797P.
XX PA (PANA-) PANACEA PHARM LLC.

PA (UYAR-) UNIV ARKANSAS.
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
PI Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals sensitive
PT to antigens, have reduced ability to bind Immunoglobulin E.
XX Disclosure; Fig 12; 100pp; English.
XX The sequence represents the amino acid sequence of modified anaphylactic
CC antigen Ara h 3, which has been altered to disrupt immunoglobulin E (IGE)
CC binding sites. Ara h 3 is an anaphylactic antigen (A), which was used to
CC design antigenic peptides having a reduced ability to bind IGE as
CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IGE binding site, where at least one IGE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IGE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IGE is reduced as compared with IGE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX SQ Sequence 526 AA;

Alignment Scores:
Pred. No.: 6,278-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 5
Query Match: 98.03% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x AAU04711 (1-526)

QY 1 CGCAGCAACCGAGGAGAGCGCGTCCAGCGCTCAATGCCAGAGACTGAC 60
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGCATTTGAATCAGAGGGCGTTTACATTGAGACTTGGAAACCCCAACCCAGAGTTC 120
Db 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnGlnGluPhe 44
QY 121 GAATCGCGCGCGTCCGCTCTCTCGTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 64
QY 181 TTCTACTCCATGCTCCCGAGAGATCTTCATCCAGAGGAGGAGGATCTTTCGGTGG 240
Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCCTGGTGTGCTAGACACTATGAAGAGCTTCACACAAGAGTTCGTCTCTCAG 300
Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
QY 301 TCCCAAAGACCAACAGAGCTCTCCAGAGAGAACCCAAAGCCAAACAGCAAGATAGT 360
Db 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
QY 361 CACCAAGAGGTGCACCGTTTCGATCAGGTGATCTCATTCAGTTCCTCCACCGCTTGTCT 420
Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144

QY	421	TTCTGGCTCTACACGACCACGACACTGATGTTGTTGCTTCTCTTACTGACACCAAC	480
Db	145	PheTrpLeuTyAenAepHisAepThrAepValValAlaValSerLeuThrAepThrAen	164
QY	481	AACAACGACACCAAGCTTGATCAGTTCCCAAGGAGATTCAATTTGGCTGGGAAACCGGAG	540
Db	165	AsnAenAepAenGlnLeuAepGlnPheProAargPheAenLeuAlaGlyAenThrGlu	184
QY	541	CAAGAGTCTTAAGGTACCAGCAACAAAGCAGACAACAAAGCAGACAGAAAGCTTTACCATAT	600
Db	185	GlnGluPheLeuAArgTyGlnGlnGlnSerArgGlnSerArgAArgAArgSerLeuProTy	204
QY	601	AGCCCATACAGCCCGCAAGTCAAGCTAGACAAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	205	SerProTySerProGlnSerGlnProAargGlnGluAargGluPheSerProAargGly	224
QY	661	CAGCACAGCCGACAGAAACGAGCAGGACAAAGAACAAACAAAGCTGTGAAACATCTTC	720
Db	225	GlnHisSerArgAArgGluAArgAlaGlyGlnGluGluAenGluGlyAenIlePhe	244
QY	721	AGCGCTTCACGCGGAGTCTCTGGAAACAAGCTTCACAGTTGACGACAGACAGATAGT	780
Db	245	SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAepAepAArgGlnIleVal	264
QY	781	CAAAACCTAGAGCGGAGACCGCAGAGTCAGAGAGGAGGCCATTGTGCACAGTCAGGGGA	840
Db	265	GlnAenLeuAargGlyGluThrGluSerGluGluGluGlyAlaIleValThrValAargGly	284
QY	841	GGCTCTCAGAATCTTTAGCCCCAGATAGAAAGAGACGTCGCCACGAAAGAGGAATACCAT	900
Db	285	GlyLeuAargAlaLeuSerProAepAArgAArgAlaAepGluGluGluTyAep	304
QY	901	GAAGATGAATATGAATACGATGAAGAGATAGAAAGCGTCGCGGGAAGCAGAGGCAGG	960
Db	305	GluAepGluTyAlaTyAepGluGluAepAArgAArgGlyAArgGlySerAArgGlyAArg	324
QY	961	GGGAATCGTATTGAAGAGACGATCTGCACCGCAGTCTAAAGAGACATTCGTAGAAC	1020
Db	325	GlyAenGlyIleGluGluThrIleCysThrAlaSerAlaIlyAlyAenIleGlyAArgAen	344
QY	1021	AGATCCCTGCACATCTCAACCCCTCAAGCTGGTTTCACTCAAAACCTGCCAACGATCTCAAC	1080
Db	345	ArgSerProAepIleTyAenProGlnAlaGlySerLeuTyThrAlaAenAepLeuAen	364
QY	1081	CTTCTAATACTTAGGTGGCTTGGACCTAGTCTGAATATGAAATCTCTACAGGAATGCA	1140
Db	365	LeuLeuIleLeuAArgTpLeuGlyProSerAlaGluTyGlyAenLeuTyAArgAenAla	384
QY	1141	TTGTTTCTCGCTCACTACACCAACGACACAGCATCATATATCATTCATTCAGGGGACGG	1200
Db	385	LeuPheValAlaHisTyAenThrAenAlaHisSerIleIleTyAArgLeuAArgGlyAArg	404
QY	1201	GCTCACGTCGAAGTCGTGGACAGCAACGGCAACAGAGTGTACACGACGAGAGCTTCAAGAG	1260
Db	405	AlaHisValGlnValValAepSerAenGlyAenAArgValTyAepGluGluLeuGlnGlu	424
QY	1261	GGTCACGTGCTGTGTGGCCACAGAACTTCGCCCTCGCTGGAAGTCCCCAGAGCAGAAC	1320
Db	425	GlyHisValLeuValValProGlnAenPheAlaValAlaGlyTySerGlnSerGluAen	444
QY	1321	TTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATACCAACCTCGCCGGTGA	1380
Db	445	PheGluTyValAlaPheTyThrAepSerAArgProSerIleAlaAenLeuAlaGlyGlu	464
QY	1381	AACTCCGCTCATAGATAACCTCGCGAGGAGGTGGTTGCCAAATTCATATGCGCTCCAAAGG	1440
Db	465	AenSerValIleAepAenLeuProGluGluValValAlaAenSerTyGlyLeuGlnAArg	484
QY	1441	GACAGGCAAGCGACGCTTAAGAACAAACACCCCTTCAAGTCTTCTCGTTCACCGTCTCAG	1500
Db	485	GluGlnAlaAargGlnLeuTyAenAenAenAenProPheTyAenPheValProProSerGln	504

QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
Db	505	GlnSerProArgAlaValAla	511
RESULT	9		
AAU05036			
ID	AAU05036	standard; protein; 526 AA.	
XX	AC	AAU05036;	
XX	XX		
XX	XX	24-OCT-2001 (first entry)	
XX	XX		
DE	XX	Modified anaphylactic peanut antigen Ara h 3.	
XX	XX		
KW	XX	Anaphylactic antigen; Ara h 3; peanut; desensitisation; antigen; allergy; immunotherapy; rush immunotherapy; anaphylaxis; asthma; food allergen;	
KW	XX	milk allergen; shellfish allergen; latex; drug; environmental allergen;	
KW	XX	grass pollen; ovalbumin; insect venom; peanut oil; immunoglobulin E; IgE.	
XX	OS	Arachis sp.	
OS	OS	Synthetic.	
XX	XX		
XX	XX	WO200139799-A2.	
XX	XX		
XX	XX	07-JUN-2001.	
XX	XX		
XX	XX	06-DEC-2000; 2000WO-US033125.	
XX	XX		
XX	XX	06-DEC-1999; 99US-00455294.	
PR	PR	23-JUN-2000; 2000US-0213765P.	
PR	PR	27-SEP-2000; 2000US-0235797P.	
XX	XX		
XX	XX	(PANA-) PANACEA PHARM LLC.	
XX	XX		
XX	XX	Caplan M;	
XX	XX		
XX	XX	WPI; 2001-408258/43.	
XX	XX		
XX	XX	Preventing allergic response against antigens, e.g. food and environmental allergens such as peanut allergen or grass pollen, latex or drug, comprises administering agent, e.g. a peptide, that blocks antigen binding sites on offending IgE.	
XX	XX		
XX	XX	Disclosure; Fig 2; 76pp; English.	
XX	XX		
CC	CC	The sequence represents the amino acid sequence of modified anaphylactic peanut antigen Ara h 3. The protein is used for active or passive desensitisation of an individual to an antigen; for alleviating or preventing allergic reactions and for decreasing the risk of allergic reactions during immunotherapy or rush immunotherapy, anaphylaxis and asthma. The antigen may be a food allergen (e.g. peanut or milk allergen), shellfish allergen, environmental allergen (e.g. grass pollen or tree pollen), latex, drug, pollen, ovalbumin, an insect venom antigen or predominantly linear epitopes. The protein is useful for protecting an individual against subsequent inadvertent or intentional exposure to antigen, e.g. receiving blocking agent before eating a chocolate bar which may inadvertently contain peanut components or before eating foods prepared using peanut oil. Administration of the blocking agents does not result in cross-linking of anti-antigenic immunoglobulin E (IgE). After exposure to the agent, the individual's antigen sensitivity is at least temporarily reduced. Only those IgE molecules that bind the offending antigen and contribute to the risk of an allergic response are blocked	
XX	XX		
SQ		Sequence 526 AA;	
Alignment Scores:			
Pred. No.:	6.27e-233	Length:	526
Score:	2638.00	Matches:	502
Percent Similarity:	99.01%	Conservative:	0
Best Local Similarity:	99.01%	Mismatches:	5
Query Match:	98.03%	Indels:	0
DB:	4	Gaps:	0

US-10-728-323-3 (1-1524) x AAU05036 (1-526)	
QY	1 CGGCAGCAACCGGAGGAGACGGTCCAGTTCACGGCTCAATCGCGAGACCTGAC 60
Db	
QY	5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
Db	
QY	61 AATCGCATGAATACAGAGGCGGTTCACATTGACACTTGGAAACCCCAACACACAGAGAGTTC 120
Db	
QY	25 AsnArgIleGlnSerGluGlyGlyIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
Db	
QY	121 GAATCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCTCGCGCGCAACGCCCTTCGTAGGCCT 180
Db	
QY	45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 64
Db	
QY	181 TTCTACTCCAATGCTCCCGAGGAGTCTTCATCCAGCAAGAGGAGGATCTTGGGTTG 240
Db	
QY	65 PheTy-SerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyPheGlyLeu 84
Db	
QY	241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAGGTGCTCGATCTCAG 300
Db	
QY	85 IlePheProGlyCysProArgHisTyGluGluProHisThrGlnGlyArgArgSerGln 104
Db	
QY	301 TCCCAAGACCACAGACTCTCCAAGAGAGAGACAAAGCCAAAGCAACGAGATAGT 360
Db	
QY	105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
Db	
QY	361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCACCGGTGTGCT 420
Db	
QY	125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
Db	
QY	421 TTCTGGCTCTACACGACCACGACACTGATGTTGTTGTTGCTTCTCTACTGACACCAAC 480
Db	
QY	145 PheTrpLeuTyRAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
Db	
QY	481 AACACGACCAACGAGCTTGATAGTTCCTCCCGAGAGATTCAATTTGCTGGGAACACGGAG 540
Db	
QY	165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
Db	
QY	541 CAAGAGTCTTAGGTACCGACCAACAAAGCAGACAAAGCAGAGAGAGAGAGTACCATAT 600
Db	
QY	185 GlnGluPheLeuArgTyRArgGlnGlnSerArgGlnSerArgArgSerLeuProTyR 204
Db	
QY	601 AGCCCATACAGCCCGCAAGTACGCTTAGCAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db	
QY	205 SerProTyR-SerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
Db	
QY	661 CAGCAGACCGCAGAGAACGAGCAGACAGAGAGAGAAACGAAAGGTGGAACATCTTC 720
Db	
QY	225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 244
Db	
QY	721 AGCGGCTTCACCGCGAGTTCCTGGAAACAGCCCTCCAGGTTACACACAGACAGATAGT 780
Db	
QY	245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264
Db	
QY	781 CAAACCTTAAGAGCGCAGACCCGAGAGTAGAGAGAGAGAGGCCATTGTGACGTGAGGGA 840
Db	
QY	265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
Db	
QY	841 GGCCTCAGAAATCTTGAGCCAGATAGAGAGACGTCGCCGACGAAGAGAGGAATACGAT 900
Db	
QY	285 GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluTyRAsp 304
Db	
QY	901 GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGGTGGCAGGGAAGCAGGAGCAGG 960
Db	
QY	305 GluAspGluTyRAlaTyRAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
Db	
QY	961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTCAAAAAGAACATTTGGTAGAAAC 1020
Db	
QY	325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleLysAsnIleGlyArgAsn 344
Db	
QY	1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCATCTCAAACTGCCAACGATCTCAAC 1080
Db	
QY	345 ArgSerProAspIleTyRAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
Db	

QY	1081 CTTCTAATACTTAGTGGCTTGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
Db	
QY	365 IeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyR-GlyAsnLeuTyRArgAsnAla 384
Db	
QY	1141 TTGTTTGTGCTCCTACTACAAACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db	
QY	385 LeuPheValAlaHisIleTyRAsnThrAsnAlaHisSerIleIleTyRArgLeuArgGlyArg 404
Db	
QY	1201 GCTCAGTCCAAGTCTGTGACAGCAACGGCAACAGAGTGTACGACGAGAGCTTCAAGAG 1260
Db	
QY	405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyRAspGluGluLeuGlnGlu 424
Db	
QY	1261 GGTCAAGTCTGTTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCACAGACGAGAAC 1320
Db	
QY	425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
Db	
QY	1321 TTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCCGCCGTGAA 1380
Db	
QY	445 PheGluTyR-ValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
Db	
QY	1381 AACTCGTTCATAGATAACCTCGCGGAGGAGGTGTTGCAAAATTCATATGGCTCCCAAGG 1440
Db	
QY	465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyR-GlyLeuGlnArg 484
Db	
QY	1441 GAGCAGGCAAGGCAGCTTAAGAACCAACACCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
Db	
QY	485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProSerGln 504
Db	
QY	1501 CAGTCTCCGAGGCTGTGGCT 1521
Db	
QY	505 GlnSerProArgAlaValAla 511
Db	
RESULT 10	
ADO38357	
ID	ADO38357 standard; protein; 507 AA.
XX	ADO38357;
XX	15-JUL-2004 (first entry)
DT	Peanut allergen Ara h 3.
DE	
XX	Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
KW	Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
KW	Immunosuppressive; Antidiabetic; Antithyroid; Antisthmatic;
KW	Antiallergic; Cytostatic; Antipsoriatic; Gene therapy; Vaccine;
KW	MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
KW	influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
KW	rickettsia; rheumatoid arthritis; multiple sclerosis;
KW	lupus erythematosus; diabetes mellitus; myasthenia gravis;
KW	autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
KW	allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
KW	adenoma; peanut; Ara h 3.
OS	Arachis hypogaea.
XX	
XX	US2004058881-A1.
XX	25-MAR-2004.
PD	
XX	24-SEP-2002; 2002US-00253286.
PF	
XX	24-SEP-2002; 2002US-00253286.
PR	
XX	(ANTI-) ANTIGEN EXPRESS INC.
PA	
XX	Humphreys RE, Xu M;
PI	
XX	WPI; 2004-294259/27.
DR	
XX	
PT	New non-naturally occurring protein or polypeptide modified by recombinant DNA techniques, useful for treating multiple sclerosis,

PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
PT colitis, cancer or psoriasis.
XX

Example 3; Page 21-22; 90pp; English.

The invention relates to a non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques comprising: a C-terminal element comprising an MHC Class II-presented epitope; an N-terminal element comprising an Ii-key motif; and an intervening element comprising a sequence of 4-11 amino acid residues where the modification by recombinant DNA techniques taking place within elements (b) and (c). Also described are methods for: suppressing or enhancing an immune response directed toward an MHC (major histocompatibility complex) Class II-presented epitope of interest. Suppressing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding an Ii-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the Ii-key motif to decrease its conformance to the archetypal Ii-Key regulatory motif. Enhancing an immune response directed toward an MHC Class II-presented epitope of interest comprises: providing a nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence encoding the MHC Class II-presented epitope of interest, the nucleic acid sequence lacking an Ii-key motif located 4-11 amino acids upstream from the N-terminal residue of the MHC Class II-presented epitope of interest; and modifying the nucleic acid sequence to introduce an Ii-key motif appropriately spaced from the MHC Class II-presented epitope. The protein or polypeptide of interest corresponds to a protein or polypeptide encoded by an infectious pathogen selected from anthrax, Ebola, HIV or influenza, preferably vaccinia virus. The non-naturally occurring protein or polypeptide (I) modified by recombinant DNA techniques is useful for treating infectious diseases caused or associated with infection by a bacterium, virus, parasite, fungus, rickettsia or other infectious agents. It is also useful for treating rheumatoid arthritis, multiple sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis, autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma, allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or adenomas. The present sequence represents the amino acid sequence of peanut allergen Ara h 3 used in the invention.

SQ Sequence 507 AA;

Alignment Scores:
Pred. No.: 8e-231 Length: 507
Score: 2615.00 Matches: 498
Percent Similarity: 98.22% Conservative: 0
Best Local Similarity: 98.22% Mismatches: 9
Query Match: 97.18% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-3 (1-1524) x AD038357 (1-507)

Qy	1	CGGCAGCAACGGAGGAGAACGGCTGTCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
Db	1	ArgGlnProGluGluAsnAlaCysGlnPheGlyArgLeuAsnAlaGlnArgProAsp	20
Qy	61	AATGCAATGAATCAGAGGCGGTTCATTGAGACTTTGGAACCCCAACACCGAGGATTC	120
Db	21	AsnArgileGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlyGluPhe	40
Qy	121	GAATGCGCGCGCTCTCTCTGCTAGTCTCCGCGCGCAACGCCCTTCGTAGGCCT	180
Db	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	60
Qy	181	TTCTACTCCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGGTG	240
Db	61	PheTyrSerAsnAlaProGlnGluIlePheIleGlyGlyArgGlyTyrPheGlyLeu	80
Qy	241	ATATTCCTGTTGCTTAGACACTATGAAGAGCCTCACACAAGGTCTGCGATCTCAG	300
Db	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlyGlyArgArgSerGln	100
Qy	301	TCCCAAGACCACCAAGACGTCTCCAGGAGAGACCAAGCCACAGCAACGAGATAGT	360

Db	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer	120
Qy	361	CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTTGCT	420
Db	121	HisGlnLysValHisArgPheAspGluGlyPheLeuIleAlaValProThrGlyValAla	140
Qy	421	TTCTGGCTCTACCAACGACGACGACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC	480
Db	141	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	160
Qy	481	AACAACGACAAACAGCTTGATTCAGTTCCCGCAGAGATTCAATTTGGTGGGAAACACGAG	540
Db	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	180
Qy	541	CAAGAGTTCTTAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600
Db	181	GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr	200
Qy	601	AGCCCATACAGCCCGCAAGTACCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
Qy	661	CAGCAGACCGCAGAGAACGAGCAGCAGCAAGAAAGAAAGAAAGGAGGAAACATCTTC	720
Db	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyAsnIlePhe	240
Qy	721	AGCGGCTTCACGCGGAGTTCTTGGAAACAGCCCTTCAGGTTGACGACAGACATAGTG	780
Db	241	SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal	260
Qy	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGA	840
Db	261	GlnAsnLeuArgGlyGluThrSerGluSerGluGluGlyAlaIleValThrValArgGly	280
Qy	841	GGCTCAGAACTTTGAGCCAGATAGAACAGAGACGTCGCGCAAGAGAGGAAATACGAT	900
Db	281	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp	300
Qy	901	GAAGATGAATATGATATCGATGAAGAGGATAGAGGCGTGGCAGGGAAGCAGAGGCGAG	960
Db	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
Qy	961	GGGAATGGTATTCAGAGACGATCTGACCGCGAAGTCTTAAAGAAACATTTGTAGAAAC	1020
Db	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
Qy	1021	AGATCCCTGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAAGATCTCAAC	1080
Db	341	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
Qy	1081	CTTCTAATACCTAGGTGGCTTGACCTAGTGTGAATATGAAATCTCTACAGGAATGCA	1140
Db	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	380
Qy	1141	TTGTTTGTGCTCCTACCAACCAACCGCACAGCATCATATATCGATTGAGGGACGG	1200
Db	381	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	400
Qy	1201	GCTCAGTCGACGTCGTGGACCAACGCAACAGAGTGTACCGCAGGAGGCTTCAAGAG	1260
Db	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	420
Qy	1261	GGTCAGCTGCTTGTGGTCACAGAACTTTCGCGCTGCTGGAAGTCCCAAGAGCGAGAAC	1320
Db	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
Qy	1321	TTCGAATAGCTGTCATTCAGACAGACTCAAGGCCAGCATATGCCAACCTCGCGGTGAA	1380
Db	441	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
Qy	1381	AACTCCGTCATAGATACCTCGCGGAGGAGTGGTTGCAATTCATATGCGCTCCAAAGG	1440

Db 461 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY 1441 GAGCAGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db 481 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCGTGTGGCT 1521
Db 501 GlnSerProArgAlaValAla 507
RESULT 11
ABG71266
ID ABG71266 standard; protein; 481 AA.
XX AC ABG71266;
XX DT 17-DEC-2002 (first entry)
XX DE Glycine max (Soybean) var. Dare protein.
XX KW Soybean; Glycinin; atomic coordinate data; processability; soya protein;
XX KM Dare; protein co-ordinate data.
XX OS Glycine max.
XX PN JP2002193996-A.
XX PD 10-JUL-2002.
XX PF 21-DEC-2000; 2000JP-00405097.
XX PR 21-DEC-2000; 2000JP-00405097.
XX PS (KYOU) UNIV KYOTO.
XX PA WPI; 2002-685438/74.
XX DR N-PSDB; ABS55193.
XX PT Glycinin, beta-conglycinin and proglycinin, their crystal structures,
XX PT three dimensional coordinates, three dimensional structured and models
XX PT and their uses.
XX PS Disclosure; Page 1273-1274; 1298pp; Japanese.
XX CC The present invention relates to a new Glycinin characterised by the
XX CC atomic coordinate data fully defined in the specification. The structure
XX CC can be used for improving processability of soya protein. The present
XX CC amino acid sequence represents the Glycine max (Soybean) var. Dare
XX CC protein, as described in the specification
XX SQ Sequence 481 AA;
Alignment Scores:
Pred. No.: 1.32e-128 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 5 Gaps: 7
US-10-728-323-3 (1-1524) x ABG71266 (1-481)
QY 1 CGGCAGCAACCGGAGGAGACGGTGCAGTTCAGCGCTCAATGCGCAGACCTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATGAATCAGAGGCGGTTACATTGACACTTGGAAACCCCAACCAACCGAGTTC 120
Db 43 AsnArgileGluSerGluGlyGlyPheileGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGCTCGCCCTCTCGCTAGTCTCGCGCGCAACCGCCCTTCGTAGCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82

QY 181 TTCTACTCCAATGCTCCCCCAGGAGATCTTTCATCCAGCAAGGAGGGGATCTTTGGGTG 240
Db 83 SerTyrThrAsnAlaProGlnGlnIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTCCTAGACACTATGAAGAGCTTCACACAAAGTTCGTGATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGly-----Gln 119
QY 301 TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGACCAAGCAACAGCAACGAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAGAGGTGCACCGTTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGTGTGTCT 420
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCTGGCTCTACAAGCACCACGACTGATGTGTGCTGTCTTCTTACTGACACCAAC 480
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACACGACAAACAGCTTCAGTTCCTCCAGGAGATTCAATTGGCTGGGACACGGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTAGACGAGAGAGCGTGAATTTAGCCCTCAGGA 660
Db 200 -----Gly 200
QY 661 CAGCAGCCGCGAGACGAGCAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGluAsnGluGlySerIleLeu 220
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAAAGCTTCCAGGTTCGACACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal--AspArgGlnIleVal 239
QY 781 CAAAACCTTAAGGCGGAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 240 ArgLysLeuGlnGlyLysGlnGluGluGluGluGluGluGluGluGluGluGluGlu 259
QY 841 GGCCTCAGATCTTCAGCCCA-----GATAGAAAGAGAGAGAGAGAGAGAGAGAG 891
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
QY 892 GAATACGATGAAGATCAATATATGATACGATGAAGAGAGATGAAGAGAGAGAGAGAG 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCAGGGGGAATGTTATGAAGAGAGAGATCTGCACCGCAAGTGTCTAAAAGAACATT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTGACATCTACACCTCAGCTGGTTCACCTAAAACCTGCCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY 1072 GATCTCAACTCTTAATCTTAGTGGCTTGGACCTTAGTGTGCAATATATGAAATCTCTAC 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGAATGCAATGTTGTGCTCCTCACTAACCAACCAAGCAGCAGCATCATATATCATG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGGACGGCTCAGCTGCAAGTCTGGGACAGCAACCGCAACAGAGGTACACAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391

Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
Qy 121 GAATCCGCGCGCTCGCTCTCTCGTTAGTCTCCGCGCAACAGCGCTTCGTAGCGCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
Qy 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATATCTTGGGTG 240
Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 102
Qy 241 ATATTCCCTGGTTCCTAGACACTATGAAGAGCTCACACAAAGGTGCTGCTATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
Qy 301 TCCCAAGAGACACAGAGCTCTCCAGGAGAGAACCAAGCCACAGCAACGAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
Qy 361 CACCAGAAAGGTGACCGTTTCATGAGGTGATCTCATTCAGTTCCACCGGTGTGCT 420
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
Qy 421 TTCTGGCTCTACACGACCGACGACACTGATGTTGCTGCTGTTCTCTTACTGACCAAC 480
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
Qy 481 AACACACACACAGCTTGATCAGTCCCGAGGAGATTCATTTGGTGGGACACGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
Qy 541 CAAGAGTTCCTAAGTACACGACCAACAGCAGCAGACAGCAGCAGCAGCAGCAGCAG 600
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 199
Qy 601 AGCCCATACAGCCCGCAAGTACAGCTAGACAGAGCGGTGAATTTAGCCCTCCAGGA 660
Db 200 -----Gly 200
Qy 661 CAGCAGAGCGCAGAGAACGAGCAGGAGCAAGAGAGAAACGAGGTGGAACATCTTC 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 220
Qy 721 AGCGGCTTCAGCGGAGTTCCTGGACAGACCTTCAGGTTGACGACAGCAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
Qy 781 CAAACCTAAGAGCGAGACCGAGAGTGAAGAGAGCGGACCATTTGACAGTGGAGGGA 840
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyLysGlyAlaIleValThrValLysGly 259
Qy 841 GGCTCAGATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAG 891
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
Qy 892 GAATACGATGAAGATGATATGATACGATGAAGAGATAGAGGCGTGGCAGGGAGAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
Qy 952 AGAGCGAGGGGGAATGTTTGAAGAGACGATCTGCACCCCAAGTGTCTAAAGAAACATT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
Qy 1012 GGTAGAACAGATCCCTGACATCTACACCTCAAGCTGTTCACTCAAAATGCCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
Qy 1072 GATCTCAACCTTCTAATCTAGTGGCTGGACCTAGTCTGCTGAATGGAATCTCTAC 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
Qy 1132 AGGAATGATGTTGTTGCTGCTACTACACCAACCAAGCAGCAGCATCATATCATGTTG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371

Qy 1192 AGGGGACGGGCTCAGCTGCAAGTCTGGAGACAGCAACGCGCAACAGAGTGTACGAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
Qy 1252 CTTCAAGAGGGTCCAGCTGCTTGTGGTCCACAGAACTTCGCCGCTCGCTGGAAGTCCCAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
Qy 1312 AGCGAGAACTTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCAACCTC 1371
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
Qy 1372 GCCGGTGAACACTCCGCTCATAGATAACCTGCCGAGAGAGTGTGCAAAATTCATATGGC 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
Qy 1432 CTCCAAGGGAGGAGCAGGCAAGCGAGCTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCA 1491
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
Qy 1492 CCCTCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db 472 ProLysGluSerGlnArgArgValValAla 481

RESULT 14

ADG43988
ID ADG43988 standard; protein; 481 AA.

AC ADG43988;

DT 26-FEB-2004 (first entry)

DE G. max glycinin G3 subunit protein.

KW oil content; plant; storage protein; seed-specific promoter; 2S-albumin;
KW 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;
KW oil production; fat production; free fatty acid production; food;
KW animal feed; pharmaceutical; fine chemical production; glycinin.

OS Glycine max.

PN WO2003077643-A2.

XX 25-SEP-2003.

XX 17-MAR-2003; 2003WO-BE002733.

XX 20-MAR-2002; 2002DE-01012893.

XX (BADI) BASF PLANT SCI GMBH.

XX Bauer J;

XX WPI; 2004-011485/01.

XX N-PSDB; ADG43987.

PT Increasing total oil content of plants, useful e.g. as foods or animal
PT feeds, by reducing amount of storage proteins, particularly with double-
PT stranded interfering RNA.

PS Claim 4; SEQ ID NO 28; 253pp; German.

XX This invention describes a novel method for increasing the total oil
CC content of a plant by reducing the amount of at least one storage protein
CC in the plant (or its tissue, organs, parts or cells) and selecting plants
CC that have higher total oil content than starting plants. The storage
CC protein is suppressed by introducing antisense RNA, optionally combined
CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding
CC factors directed against storage protein genes, viral sequences that
CC degrade storage protein RNA, constructs that induce homologous
CC recombination of endogenous storage protein genes or mutations into
CC storage protein genes. Most preferably a plant cell is stably transfected

CC with a recombinant expression construct, then regenerated to plants that
 CC express the incorporated sequence. The expression constructs particularly
 CC contain a seed-specific promoter and they are introduced into plants by
 CC standard methods, e.g. via Agrobacterium. The preferred storage proteins
 CC of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-
 CC prolamines. Transgenic organisms produced by the new method are used for
 CC production of oils, fats, free fatty acids or their derivatives, useful
 CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence
 CC represents a storage protein used to illustrate the method of the
 CC invention.

XX SQ Sequence 481 AA;

Alignment Scores:

Pred. No.: 1 32e-128 Length: 481
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.57% Conservative: 68
 Best Local Similarity: 58.24% Mismatches: 91
 Query Match: 55.82% Indels: 54
 DB: 8 Gaps: 7

US-10-728-323-3 (1-1524) x ADG43988 (1-481)

QY 1 CGCAGCAACCGGAGGAGACGGCTCCAGTTCAGCGCTCAATGCGCAGACCTGAC 60
 DB 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
 QY 61 AATCGCATTAATCAGAGGCGGTACATTGACACTTGGNACCCCAACACGAGGATTC 120
 DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
 QY 121 GAATGGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
 DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
 QY 181 TTCTACTCCATCTCCCCAGGAGATCTTCATCCAGCAAGGAGGATATTTGGGTG 240
 DB 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 102
 QY 241 ATATTCCTCGTGTCTGACACTATGAGAGCTTCACACAAAGTCTGTCATCTCAG 300
 DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
 QY 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGAACCAAGCCCAACAGCAACGAGATAGT 360
 DB 120 SerSerArgPro-----GlnAspArg 126
 QY 361 CACCAAGGTGACCGTTTCGATGAGGTGATCTCATTCAGTTCACCGGTGTGCT 420
 DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
 QY 421 TTCTGGCTCTACACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCCMAC 480
 DB 147 TyrTrpMetTyrAsnAsnGluAsnGluAsnThrProValAlaValSerLeuIleAspThrAsn 166
 QY 481 AACACGACACCAACGAGTTGATTCCTCCCGCAGAGATTCATTTGGTGGGAACGAG 540
 DB 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrIleAlaGlyAsnGlnGlu 186
 QY 541 CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
 DB 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 199
 QY 601 AGCCCATACCGCGGNAAGTCAGCTAGACAGAGAGCGTGAATTTAGCCCTCGAGGA 660
 DB 200 -----Gly 200
 QY 661 CAGCAGACGCGCAGAGAAACGAGCAGGACAGAGAAAGAAACGAAAGTGGAAACATCTTC 720
 DB 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGlnGluAsnGluGlySerIleLeu 220
 QY 721 AGCGGCTTACGCGGAGTTCTCTGGAAACAAGCCCTTCAGGTTGACGACAGACAGATAGT 780

DB 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
 QY 781 CAAAACCTAAGAGGCGGAGACCGAGAGTGAAGAGAGGAGGAGCCATTCTGACAGTGGGGA 840
 DB 240 ArgLysLeuGlnGlnGluAsnGluGluGluLysGlyAlaIleValThrValLysGly 259
 QY 841 GGCCTCAGAACTTTCAGCCCA-----GATAGAAAGAGACGTGCCGACGAAAGAG 891
 DB 260 GlyLeuSerValIleSerProThrGluGlnGlnArgProGluGluGluGlu 279
 QY 892 GAATACGATGAAGATGAATATGATACGATGAAGAGGATAGAGGCGTGGCAGGGAAGC 951
 DB 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
 QY 952 AGAGGCGGCGGATGTTTGAAGAGACGATCTGCACCGCAAGTCTCTAAAAGACATT 1011
 DB 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
 QY 1012 GGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGTTCACATAAACTGCAAC 1071
 DB 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
 QY 1072 GATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTCTGCTGAATATGAAATCTCTAC 1131
 DB 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
 QY 1132 AGGAATGCAATTGTTGCTCCTCACTACACCAACCGCAGCAGACGATGTACCGAGGAG 1191
 DB 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
 QY 1192 AGGGGACGGGCTCACGTGCAAGTCTGTGAGCAGCAGCAGCAGCAGTGTACCGAGGAG 1251
 DB 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
 QY 1252 CTTCAAGAGGTCACGTGCTTGTGTCGCCACAGAACTTCGCGTCCGTGGAAGTCCCAG 1311
 DB 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
 QY 1312 ACGGAACTTCGAATACGTGCTTCAAGACAGACTCAAGGCCAGCAGCAGTGTACCAACCTC 1371
 DB 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
 QY 1372 GCCGTGAAAACTCCGTTCATAGATACTCCCGAGGAGGTGTTGCCAAATTCATATGC 1431
 DB 432 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
 QY 1432 CTTCAAGGAGCAGCAGCAGCAGCTTAAGAACACAAACCCCTTCAAGTCTTCGTCCA 1491
 DB 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
 QY 1492 CGGTCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
 DB 472 ProLysGluSerGlnArgArgValValAla 481
 RESULT 15
 ABG71265
 ID ABG71265 standard; protein; 485 AA.
 XX
 AC ABG71265;
 XX
 DT 17-DEC-2002 (first entry)
 XX
 DE Glycine max (Soybean) var. Shiotsurunoko protein #2.
 XX
 XX Soybean; Glycinin; atomic coordinate data; processability; soya protein;
 XX Shiotsurunoko; protein co-ordinate data.
 XX
 OS Glycine max.
 XX
 PN JP2002193996-A.
 XX
 XX 10-JUL-2002.
 PD

PF 21-DEC-2000; 2000JP-00405097.
 PR 21-DEC-2000; 2000JP-00405097.
 XX (KYOU) UNIV KYOTO.
 XX WPI; 2002-685438/74.
 DR N-PSDB; ABS55192.
 XX Glycine, beta-conglycinin and proglycinin, their crystal structures,
 PT three dimensional coordinates, three dimensional structured and models
 PT and their uses.
 XX Disclosure; Page 1269-1271; 1298pp; Japanese.
 XX The present invention relates to a new Glycinin characterised by the
 CC atomic coordinate data fully defined in the specification. The structure
 CC can be used for improving processability of soya protein. The present
 CC amino acid sequence represents the Glycine max (Soybean) var.
 CC Shirotaurunoko protein #2, as described in the specification
 XX
 SQ Sequence 485 AA;
 Alignment Scores:
 Pred. No.: 2,25e-128 Length: 485
 Score: 1499.50 Matches: 297
 Percent Similarity: 72.02% Conservative: 71
 Best Local Similarity: 58.12% Mismatches: 94
 Query Match: 55.72% Indels: 49
 DB: 5 Gaps: 6
 US-10-728-323-3 (1-1524) x ABG71265 (1-485)
 QY 1 CGG CAG CAA CGG A G A G A G A G C G T G C C A G T T C C A G C C C T C A A T G C G C A G A C C T G A C 60
 DB 20 A R G G L u G l n A l a G l n A s n G l u C y s G l n l e g l n l y s L e u A s n A l a L e u L y s P r o A s p 39
 QY 61 A A T C G C A T T G A A C A G A G G G C G T T A C A T T G A G A C T T G G A C C C C A C A C C A G G A G T T C 120
 DB 40 A s n A r g l e G l u s e r G l u G l y P h e i l e G l u t h r T r p A s n P r o A s n A s n L y s P r o P h e 59
 QY 121 G A A T G C C G C G C G C C T C T C T C G C T T A G T C T C C G C G C A A C G G C C T C G T A G G C C T 180
 DB 60 G l n C y s A l a G l y V a l A l a L e u S e r A r g C y s T h r L e u A s n A r g A s n A l a L e u A r g P r o 79
 QY 181 T T C T A C C A A T G C T C C C A G G A G A T C T T C A T C A G C A A G A A G G G A T A C T T T G G G T T G 240
 DB 80 S e r T y r T h r A s n G l y P r o G l n G l u I l e T y r I l e G l n G l n G l y A s n G l y I l e P h e G l y M e t 99
 QY 241 A T A T T C C C T G T T G T C T A G A C A C T A T G A A G A G C C T C A C A C A G A G T C G T C G A T C T C A G 300
 DB 100 I l e P h e P r o G l y C y s P r o S e r T h r T y r G l n G l u P r o G l n G l u S e r G l n G l n A r G g l y A r g 119
 QY 301 T C C C A A A G A C C A C A A G A C G T C T C C A A G G A G A A G A C C A A G C C A C A G C A A C A G A T A G T 360
 DB 120 S e r G l n A r g P r o G l n A s p A r g 126
 QY 361 C A C C A G A A G T G C A C C G T T T C G A T G A G G G T G A T C T A T T G C A G T T C C A C C G G T G T T G C T 420
 DB 127 H i s G l n l y s V a l H i s A r g P h e A r g G l u G l y A s p L e u i l e A l a V a l P r o T h r G l y V a l A l a 146
 QY 421 T T C T G G T C T A C A A C G A C C A G A C A C T G A T G T T G T T G T T C T T C T T A C T G A C C A A C 480
 DB 147 T r p T r p M e t T y r A s n A s n G l u A s p T h r P r o V a l V a l A l a V a l S e r I l e A s p T h r A s n 166
 QY 481 A A C A A C A C A C A C C G T T G A T C A G T T C C C A G G A G A T T C A A T T G C T G G G A C A C G G A G 540
 DB 167 S e r L e u G l n A s n G l n L e u A s p G l n M e t P r o A r g A r g P h e T y r L e u A l a G l y A s n G l n G l u 186
 QY 541 C A A G A G T T C T T A A G G T A C C A G C A A C A A G C A G A C A A A G C A G A C G A A G A A G C T T A C C A T A T 600
 DB 187 G l n G l u P h e L e u l y s T y r G l n G l n G l n ----- 195

QY 601 A G C C C A T A C A G C C C G C A A A G T C A G C C T A G A C A A G A G C G T G A A T T T A G C C C T C G A G G A 660
 DB 196 ----- G l n G l n G l y G l y S e r G l n S e r G l n L y s G l y 205
 QY 661 C A G C A C A C C C G C A G A A C G A G C A G A G A A A A C G A A A G T G A A A C A C T C T T C 720
 DB 206 L y s G l n ----- G l n G l u G l u A s n G l u G l y S e r A s n I l e L e u 218
 QY 721 A G C G G C T T C A C G C G G A G T T C T G G A A C A A G C C T T C A G T T C A C C A C A C A G A T A G T G 780
 DB 219 S e r G l y P h e A l a P r o G l u P h e L e u L y s G l u A l a P h e G l y V a l --- A s n M e t G l n I l e V a l 237
 QY 781 C A A A C C T A A G A G C G A G C C G A G A G T G A A G A G A G A G C C A T T G T G A C A G T A G C G G A 840
 DB 238 A r g A s n L e u G l n G l y G l u A s n G l u G l u A s p S e r G l y A l a I l e V a l T h r V a l l y s G l y 257
 QY 841 G G C C T C A G A A T C T T G A C C C A G A T A G A A A G A C G T G C C G A C C A A G A A G A G A A T A C G A T 900
 DB 258 G l y L e u A r g V a l T h r A l a P r o A l a M e t A r g L y s P r o G l n G l n G l u A s p A s p A s p 277
 QY 901 G A A G A T G A A ----- T A T G A A T A C G A T G A A G A G A T A G A A G G C T G G C A G G G A 948
 DB 278 G l u G l u G l n P r o G l n C y s V a l G l u T h r A s p L y s G l y C y s G l n A r g G l n S e r L y s A r g 297
 QY 949 A G C A G A G G C A G G G G A A T G T A T T A A G A G A C A G A T C T G C A C C C A A G T G C T A A A A A G A A C 1008
 DB 298 S e r A r g ----- A s n G l y I l e A s p G l u T h r I l e C y s T h r M e t A r g L e u A r g G l n A s n 314
 QY 1009 A T T G G T A G A A A C A G A T C C C C T G A C A T C T A C A A C C C T C A A G C T G T T C A C T C A A A A C T G C C 1068
 DB 315 I l e G l y G l n A s n S e r S e r P r o A s p I l e T y r A s n P r o G l n A l a G l y S e r I l e T h r A l a 334
 QY 1069 A A G A T C T C A A C C T T C T A T A C T A G T G C T T G G A C C T A G T G C T G A A T A T G A A A T C T C 1128
 DB 335 T h r S e r L e u A s p P h e P r o A l a L e u T p L e u L e u L y s L e u S e r A l a G l n T y r G l y S e r L e u 354
 QY 1129 T A C A G A A T G C A T T G T T C T C A C T A C A C A C C A C C A G C A C A C A G C A T C A T A T A T C G A 1188
 DB 355 A r g L y A s n A l a M e t V a l P r o H i s T y r T h r L e u A s n A l a A s n S e r I l e I l e T y r A l a 374
 QY 1189 T T C A G G G A C G G G C T C A C G T G A A G T C G T G G A C A G C A A C G G C A A C A G A G T G T A C G A C G A G 1248
 DB 375 L e u A s n G l y A r g A l a L e u V a l G l n V a l A s n C y s A s n G l y G l u A r g V a l P h e A s p G l y 394
 QY 1249 G A G C T T C A A G A G G T C A C G T G T T G T G T G C C A C A G A A C T T C G C C T C G C T G A A A G T C C 1308
 DB 395 G l u L e u G l n G l u G l y V a l L e u I l e V a l P r o G l n A s n P h e A l a V a l A l a L y s S e r 414
 QY 1309 C A G A G C G A A C T T C G A A T A C G T G G C A T T C A A G A C A G A C T C A A G C C C A G C A T A G C C A A C 1368
 DB 415 G l n S e r A s p A s n P h e G l u T y r V a l S e r P h e L y s T h r A s n A s p A r g P r o S e r I l e G l y A s n 434
 QY 1369 C T C G C G G T G A A A A C T C C G T C A T A G A T A A C C T G C C G A G A G G T G T T G C A A A T T C A T A T 1428
 DB 435 L e u A l a G l y A l a A s n S e r L e u L e u A s n A l a L e u P r o G l u G l u V a l l e G l n H i s T h r P h e 454
 QY 1429 G G C C T C C A A G G A G C A G C A A G C A G T T A A C A C A C A C C C C T T C A A G T T C T C G T T 1488
 DB 455 A s n L e u L y s S e r G l n G l n A l a A r g G l n V a l L y s A s n A s n A s n P r o P h e S e r P h e L e u V a l 474
 QY 1489 C A C C G T C T C A G C A G T C T C G A G G G C T G T G G C T 1521
 DB 475 P r o P r o G l n G l u S e r G l n A r g A r g A l a V a l A l a 485

Search completed: August 24, 2005, 09:56:21
 Job time : 110.311 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 07:15:51 ; Search time 19.0812 Seconds
(without alignments)
15369.504 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 cggcagcaaccgagagaa.....ctccgagggtgtggttaa 1524

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10728323/runat 23082005 124355 29222/app query.fasta_1.4757
-DB=pir 79 -QFMT=fastcan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPT=0 -LOOPT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10728323 @CGN 1.1 76 @runat 23082005 124355 29222 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1502	55.8	481	2 S04605	glycinin G3 - soyb
2	1499.5	55.7	485	2 S11002	glycinin G2 precu
3	1496	55.6	495	2 S10851	glycinin G1 precu
4	1492.5	55.5	485	1 FWSYG1	glycinin chain A2B
5	1466	54.5	495	1 FWSYG2	glycinin chain Ala
6	1457.5	54.2	498	2 S44294	legumin A precursor
7	1455	54.1	520	2 S08237	legumin A2 precu
8	1440	53.5	517	1 FWPMLA	legumin A precursor
9	1431.5	53.2	482	2 S49877	legumin A precursor
10	1428	53.1	484	2 S11003	legumin G3 precu
11	1414.5	52.6	500	2 S14393	legumin A2 precu
12	1384	51.4	507	2 T06452	probable legumin A
13	1375	51.1	497	2 S14392	legumin A1 precu
14	1073.5	39.9	551	2 S51941	prunin 1 precursor

15	1049	39.0	560	2 S11004	glycinin G4 precu
16	1032	38.4	503	2 S00336	legumin B legu pre
17	1019.5	37.9	500	2 S26688	legumin K - garden
18	999.5	37.1	566	2 T06453	probable legumin B
19	998.5	37.1	484	2 A24942	legumin B4 precu
20	998	37.1	485	2 S44268	legumin B precursor
21	981.5	36.5	564	2 S37241	legumin B - fava b
22	974	36.2	563	2 S54802	glycinin ASA4B3 ch
23	973.5	36.2	562	1 FWSYG5	glycinin chain ASA
24	968	36.0	504	2 S51942	prunin 2 precursor
25	946.5	35.2	562	2 S20946	glycinin Gv4 precu
26	941.5	35.0	507	1 FWCNBA	beta-globulin A pr
27	926.5	34.4	516	1 FWSYG3	glycinin G5 precu
28	916	34.0	516	1 FWCNBB	beta-globulin B pr
29	911	33.9	470	2 S68964	legumin precursor
30	910	33.8	476	2 S68965	legumin precursor
31	892.5	33.2	501	2 S49422	11S globulin seed
32	873.5	32.5	472	2 S68966	legumin precursor
33	846.5	31.5	480	1 FWPULB	11S globulin beta
34	844.5	31.4	499	2 S06350	glutelin type 1 pr
35	841.5	31.3	499	2 B34332	glutelin precursor
36	837.5	31.1	499	2 A27033	glutelin 2 precu
37	837	31.1	499	2 S17762	glutelin gluB-1 pr
38	835	31.0	499	2 S07640	glutelin precursor
39	832	30.9	315	2 JC2094	legumin type A alp
40	830.5	30.9	499	2 A34332	glutelin II precu
41	829.5	30.8	499	1 FWR22	glutelin II precu
42	827.5	30.8	500	2 S05443	glutelin precursor
43	824.5	30.6	496	2 S18745	glutelin - rice
44	822.5	30.6	307	2 PC2050	legumin type A alp
45	820	30.5	496	2 S17763	glutelin gluB-2 pr

ALIGNMENTS

RESULT 1
S04605
Glycinin G3 - soybean
C;Species: Glycine max (soybean)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S04605
R;Cho, T.J.; Nielsen, N.C.
Nucleic Acids Res. 17, 4388, 1989
A;Title: The glycinin Gy(3) gene from soybean.
A;Reference number: S04605; MUID:89296500; PMID:2740231
A;Accession: S04605
A;Molecule type: DNA
A;Residues: 1-481 <CHO>
A;Cross-references: UNIPROT:P11828; EMBL:X15123; NID:g18638; PIDN:CAA33217.1; PID:g18639
A;Experimental source: variety Dare
C;Genetics:
A;Gene: Gy3
A;Introns: 96/1; 177/3; 352/3
C;Superfamily: glycinin

Alignment Scores:
Pred. No.: 1.02e-103 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 2 Gaps: 7

US-10-728-323-3 (1-1524) x S04605 (1-481)

Qy	1	CGCAGCACCCGAGGAGACCGCGTCCAGTTCACGCGCTCAATCGCAGAGACCTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuIlysProAsp	42
Qy	61	AATCGATTGAATCAGAGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnIlysProPhe	62

Qy	1	CGGCAGCAACCGGAGAGAACCGCTGTCGAGTTCGAGCGCTCAATGCGCAGAGACCTGAC	60
Db	20	ArgGluGlnAlaGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp	39
Qy	61	AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAACTGGACCCCAACACAGGAGTTC	120
Db	40	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	59
Qy	121	GAATCGCCGGCGTGGCGCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGGCGCT	180
Db	60	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro	79
Qy	181	TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGTTG	240
Db	80	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet	99
Qy	241	ATATTCCCTGGTGTCTCTAGACACTATGAGAGCCTCACACAAAGTCTCGTCACTCAG	300
Db	100	IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg	119
Qy	301	TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGAACCAAGACCAACAGCAACAGATAGT	360
Db	120	SerGlnArgPro-----GlnAspArg	126
Qy	361	CACCAGAAGTGCA CCGTTTCGATGAGGGTCAATCTCATTGAGTTCACCGGTGTGCT	420
Db	127	HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	146
Qy	421	TTCTGGCTCTACACGACGACGACACATGATGTTGCTGCTGTTCTCTTACTGACACCAAC	480
Db	147	TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn	166
Qy	481	AACAACGACCAACGACTTGATTCAGTCTCCAGGAGATTCAATTGGCTGGGAACACGGAG	540
Db	167	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
Qy	541	CAAGAGTCTTAAAGGTACCAACGACAAAGCAGACAAAAGCAGACGAAAGACTTACCATAT	600
Db	187	GlnGluPheLeuLysTyrGlnGln-----GlnGlnGlySerGlnSerGlnLysGly	205
Qy	601	AGCCCATACAGCCCGCAAGCTAGACGAGAGCGTGATTTAGCCCTCCAGGA	660
Db	196	-----GlnGlnGlySerGlnSerGlnLysGly	205
Qy	661	CAGCACAGCCGACAGAAACGACGAGGACAAAGAAAGAAACGAGGTGAAACATCTTC	720
Db	206	LysGln-----GlnGlnGluAsnGluGlySerAsnIleLeu	218
Qy	721	AGCGGCTTCAGCCGGAGTCTCTGGAAACAAGCTTCAGGTTCACGACACAGATAGTG	780
Db	219	SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal	237
Qy	781	CAAACTTACAGCGGACGAGAGTGAAGAGAGGGGCCATTGTGACAGTGAAGGGGA	840
Db	238	ArgAsnLeuGlnGlyGluAsnGluGluGluAspSerGlyAlaIleValThrValLysGly	257
Qy	841	GGCCTCAGATCTTGAGCCCGACATAGAAAGAGAGCTCCCGACGAGAGAGGAAATACGAT	900
Db	258	GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAlaAspAspAsp	277
Qy	901	GAAGATGAA-----TATGAATACGATGAAGAGGATGAAGCGCTGGCAGGGGA	948
Db	278	GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg	297
Qy	949	AGCAGAGCGAGGGGAATGTGATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGAAC	1008
Db	298	SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn	314
Qy	1009	ATTGGTAGAAACAGATCCCTCGACATCTACACCTCAAGCTGGTTCACATCAAACTGCC	1068
Db	315	IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrThrAla	334

Qy	1069	AACGATCTCAACCTTCTAATACCTTAGTGGCTTGGACCTAGTCTGCTGAATATGAAATCTC	1121
Db	335	ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTrpGlySerLeu	354
Qy	1129	TACAGGAATGCATTGTTGTTCGCTCACTACAAACCAACCAACGACACAGCATCATATATCGA	1188
Db	355	ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla	374
Qy	1189	TTGAGGGACGGGCTCACTGTCGAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAG	1248
Db	375	LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly	394
Qy	1249	GAGCTTCAAGAGGGTCACGCTGCTTGGTCCACAGAACTTCGCCGCTCGCTGGAAGTCC	1308
Db	395	GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaAlaLysSer	414
Qy	1309	CAGAGCAGAACTTCGAATACGTGGCATTCAAGACAGACTCAAGGCCACCATAGCCCAAC	1368
Db	415	GlnSerAspAsnPheGluTrpValSerPheLysThrAsnAspArgProSerIleGlyAsn	434
Qy	1369	CTGCCCGGTGAAAACTCCGTTCATGATAAACCCTCCGGAGGAGGTGGTTGCAAAATTCATAT	1428
Db	435	LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe	454
Qy	1429	GGCTTCCAAAGGGACGAGCAAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTT	1488
Db	455	AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal	474
Qy	1489	CCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT	1521
Db	475	ProProGlnGluSerGlnArgAlaValAla	485
RESULT 3			
S10851			
glycinin G1 precursor - soybean			
N:Alternate names: Glycinin AlaBx			
N:Contains: Glycinin chain Ala; glycinin chain Bx			
C:Species: Glycine max (soybean)			
C>Date: 21-Nov-1993 #sequence revision 19-Jan-1996 #text_change 09-Jul-2004			
C:Accession: S10851; S04603; JS0015			
P:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fische			
Plant Cell 1, 313-328, 1989			
A:Title: Characterization of the glycinin gene family in soybean.			
A:Reference number: S10851; MUID:92393391; PMID:2485233			
A:Accession: S10851			
A:Molecule type: DNA			
A:Residues: 1-495 <NIE>			
A:Cross-references: UNIPROT:P04776			
A:Experimental source: variety Dare			
R:Sims, T.L.; Goldberg, R.B.			
Nucleic Acids Res. 17, 4386, 1989			
A:Title: The glycinin Gy(1) gene from soybean.			
A:Reference number: S04603; MUID:89296498; PMID:2740229			
A:Accession: S04603			
A:Molecule type: DNA			
A:Residues: 1-495 <SIM>			
A:Cross-references: EMBL:X15121; NID:g18634; PIDN:CAA33215.1; PID:g18635			
A:Experimental source: variety Dare			
R:Utsunomi, S.; Kohno, M.; Mori, T.; Kito, M.			
J. Agric. Food Chem. 35, 210-214, 1987			
A:Title: An alternate cDNA encoding glycinin Ala Bx subunit.			
A:Reference number: JS0015			
A:Accession: JS0015			
A:Molecule type: mRNA			
A:Residues: 1-495 <UTS>			
A:Experimental source: var. Shirotsurunoko			
C:Genetics:			
A:Gene: Gy1			
A:Introns: 96/1; 180/3; 366/3			
C:Superfamily: Glycinin			
C:Keywords: storage protein			
F:1-19/Domain: signal sequence #status predicted <SIG>			
F:20-495/Product: glycinin G1 #status predicted <MATL>			

F;2-310/Product: glycinin Ala chain #status predicted <MAT2>
F;311-495/Product: glycinin Bx chain #status predicted <MAT3>

Alignment Scores:

Pred. No.: 2,87e-103 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.87% Conservative: 73
Best Local Similarity: 56.70% Mismatches: 100
Query Match: 55.59% Indels: 50
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x S10851 (1-495)

```
QY 1 CGGCAGCAACCGGAGACCGCTGCCAGTTCACAGCGCTCAATGCGCAGACCTGAC 60
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 42
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AATCGCATGAATCAGAGCGCGTTACATTGACACTTGGAAACCCCAACACCGAGATTTC 120
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 43 AsnArgileGluSerGluGlyGlyLeuIleGluThrTriAsnProAsnAsnLysProPhe 62
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 GAATGGCGCGCTCGCCCTCTCTCGCTAGTCTCTCGCGCGCAACGCCCTTCGTAGCGCT 180
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 TTCTACTCCAACTGCCCGCAGGAGATCTTCATCCAGCAAGGAAGGGATACTTTGGGTTG 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGACCTCACACAAAGGTTCGTGATCTCAG 300
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 IleTyrProGlyCysProSerThrPheGluGluProGlnGlnArgGlyGln 122
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 TCCCAAGACCACCAAGACTCTCCAAGGAGAGACCAAGCCCAACAGCAACAGATAGT 360
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 123 SerSerArgPro-----GlnAspArg 129
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 CACGAGAGGTGCACGGTTTCGATAGGGTGATCTCATTGCAGTTCGCCACCGGTGTGCT 420
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 130 HisGlnLysIleTyrAsnPheArgGluGlyAspIleAlaValProThrGlyValAla 149
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 TTCTGCTCTACACGACACACGACACTGATGTTGTTCTCTCTACTGACACCAAC 480
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 150 TriPTriMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn 169
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 AACACGACACACGACTTGATGATCTCCCGCAGGAGATTCAATTGCTGGGAACACGGAG 540
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 170 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 541 CAAGAGTTCTTAAGGTACACGACCAACAAAGCAGACAAAGCAGACGAAAGACTTACCATAT 600
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 190 GlnGluPheLeuLysTyrGlnGlnGluGln----- 199
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTAGCCTAGACAAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 200 -----GlyGly 201
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 661 CACGACAGCCGACAGNACGACGAGCAGCAGACAGAGAAAGCAAGCTGGAACATCTTC 720
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 202 HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlyGlySerIleLeu 221
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 721 AGCGGCTTCCAGCGCGGAGTTCCTGGAAACAGCCCTTCAGGTTCCACGACAGACAGATAGTG 780
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 222 SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla 240
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 781 CAAAACTAAGAGCGGACCGCAGAGAGTAGAAGAGAGAGAGCGCCATTGTGACGTAGGGGA 840
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly 260
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 841 GGCCTCAGATCTTGACCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAG 891
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 261 GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu 280
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 892 GAATACGATGAAGATGAATATGAATAC-----GATCAAGC---GATAGAAGG 936
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 937 COTGCGAGGGGAAGCAGAGCGAGGGGAATGTTATTGAAGACCATCTGCGCCGCAAGT 996
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 ArgGlySerGlnSerLysSerArgAsnGlyIleAspGluThrIleCysThrMetArg 320
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 997 GCTAAAGAACATTCGTAGAAACAGATCCCTGCATCTACAAACCTCAACCTGGTTCA 1056
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 321 LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer 340
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1057 CTCAAAACTGCCAACCATCTCAACCTTCTAATACTTAGTGCTTGGAGACTAGTGTGAA 1116
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 341 ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTriPLeuArgLeuSerAlaGlu 360
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1117 TATGGAATCTCTACAGGAATGCANTGTTTGTGCTCTACTACAACACCAACGACACAGC 1176
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer 380
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1177 ATCATATATCGATTGAGGGGACGGCTCACGTGCAAGTCGTGACAGCAACCGCAACAGA 1236
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 381 IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValAlaAsnCysAsnGlyGluArg 400
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1237 GTGTACGACGAGAGCTTCAAGAGGCTCACGTGCTTGTGTCGCCACAGAACTTCGCCGTC 1296
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 401 ValPheAspGlyGluLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValVal 420
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1297 GCTGGAAGTCCGAGGAGCAACTTCGAATACGTGGGATTCGAAGACAGACTCAAGGCC 1356
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 AlaAlaArgSerGlnSerAspAsnGluTyrValSerPheLysThrAsnAspThrPro 440
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1357 AGCATAGCCAACTCGCCGTCGAACTCCGTCATAGATAACCTCCGAGGAGGTGGTT 1416
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 441 MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluValIle 460
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1417 GCAATTCATATGCGCTCCAAAGGAGGAGGAGGCAAGCGAGCTTAAAGAACAAACCCCTTC 1476
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 461 GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnProPhe 480
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1477 AAGTTCCTTCGTTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 LysPheLeuValProGlnGluSerGlnLysArgAlaValAla 495
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 4
PWSYGI
glycinin chain A2Bla precursor - soybean
N;Alternate names: 11S globulin
N;Contains: glycinin chain A2; glycinin chain Bla
C;Species: Glycine max (soybean)
C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 09-Jul-2004
C;Accession: A91341; A92454; E92454; A90024; A92452; S10503; S74123; A05082; A05164; A051
R;Momma, T.; Negro, T.; Uda, K.; Fukazawa, C.
FEBs letc. 188, 117-122, 1985
A;Title: A complete cDNA coding for the sequence of glycinin A2Bla subunit precursor.
A;Reference number: A91341
A;Accession: A91341
A;Molecule type: mRNA
A;Residues: 1-485 <MOM>
A;Cross-references: UNIPROT:P04405
A;Experimental source: strain Bonimori
A;Note: the source of this mRNA was cotyledon tissue taken from seeds at the middle stage
R;Marco, Y.A.; Thanh, V.H.; Tumer, N.E.; Scallion, B.J.; Nielsen, N.C.
J. Biol. Chem. 259, 13436-13441, 1984
A;Title: Cloning and structural analysis of DNA encoding an A2Bla subunit of glycinin.
A;Reference number: A92454; MUID:85030472; PMID:6092376
A;Accession: A92454
A;Molecule type: mRNA
A;Residues: 262-446 <MA1>
A;Accession: B92454
A;Molecule type: DNA
A;Residues: 318-485 <MA2>
A;Experimental source: strain CX635-1-1-1
```


QY	1249	GAGCTTCAAGAGGTCACGTCCTTGCTGGTCCACAGAACTTCGCCGCTCGCTGGAAAGTCC	1308
Db	395	GlueGlnGluGlyGlyValLeuLeuValProGlnAsnPheAlaValAlaLysSer	414
QY	1309	CAGAGCAGAACTTCGATACGTGGCATTTCAAGACAGACTCAAGGCCACCATAGCCAAC	1368
Db	415	GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn	434
QY	1369	CTGCCCGTCAAACTCCGTCATAGATAACCTCGCGAGGAGGTGGTTCAAAATTCATAT	1428
Db	435	LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluValIleGlnHisThrPhe	454
QY	1429	GGCTTCAAGAGGAGCAGGCAAGCAGCTTAAAGCAACACCCCTTCAAGTCTTCGTT	1488
Db	455	AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal	474
QY	1489	CCACCGTCTCAGCAGTCTCCGAGGGCTGTGCT	1521
Db	475	ProProGlnGluSerGlnArgArgAlaValAla	485
RESULT 5			
FWSYG2			
glycinin chain AlaBx precursor - soybean			
N;Alternate names: IIS globulin; glycinin AlaBib			
C;Species: Glycine max (soybean)			
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004			
C;Accession: A23497; S10502			
R;Negoro, T.; Momma, T.; Fukazawa, C.			
Nucleic Acids Res. 13, 6713-6731, 1985			
A;Title: A cDNA clone encoding a glycinin A-1a subunit precursor of soybean.			
A;Reference number: A23497; MUID:86041867; PMID:2997720			
A;Accession: A23497			
A;Molecule type: mRNA			
A;Residues: 1-495 <NEG>			
A;Cross-references: UNIPROT:P04776; GB:X02985; NID:g18614; PIDN:CAA26723.1; PID:g18615			
A;Experimental source: cv. Boninori			
A;Note: the authors translated the codon AAC for residue 449 as Lys			
R;Kitamura, Y.; Arahira, M.; Itoh, Y.; Fukazawa, C.			
Nucleic Acids Res. 18, 4245, 1990			
A;Title: The complete nucleotide sequence of soybean glycinin A2B1a gene spanning to and			
A;Reference number: S10502; MUID:90332420; PMID:2377465			
A;Accession: S10502			
A;Status: preliminary; translation not shown			
A;Molecule type: DNA			
A;Residues: 481-495 <KIT>			
A;Cross-references: EMBL:X53404; NID:g18522; PIDN:CAA37479.1; PID:g18523			
C;Comment: The source of this protein was cotyledon tissue taken 38 days after flowering			
C;Comment: By homology with the A2B1a component, residues 307-310, and 491-495 are removed			
C;Superfamily: glycinin			
C;Keywords: seed; storage protein			
F;1-19/Domain: signal sequence #status predicted <SIG>			
F;20-306/Product: glycinin chain Ala #status predicted <CLA>			
F;311-490/Product: glycinin chain Bx #status predicted <GUB>			
F;107-317/Disulfide bonds: #status predicted			
Alignment Scores:			
Pred. No.:	4, 88e-101	Length:	495
Score:	1466.00	Matches:	288
Percent Similarity:	70.10%	Conservative:	73
Best Local Similarity:	55.92%	Mismatches:	104
Query Match:	54.48%	Indels:	50
DB:	1	Gaps:	6
US-10-728-323-3 (1-1524) x FWSYG2 (1-495)			
QY	1	CGCAGCAACCGGAGGAGAACGGCTGCCAGTTCAGCGCTCAATCGCGAGACCTGAC	60
Db	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProGly	42
QY	61	AATCGCATTAATCAGAGGGCGGTATCATTTGACACTTGGACCCCAACACACGAGGATTC	120
Db	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe	62

QY	121	GAATGCCGCGGCTCGCCCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCCT	180
Db	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
QY	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAAGGATCTTTGGGTG	240
Db	83	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet	102
QY	241	ATATTCCCTGGTTCCTAGACACTATGAAGAGCCTCACACAAAGGTGCTCGATCTCAG	300
Db	103	IleTyrProGlyCysSerSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln	122
QY	301	TTCCAAAGACCCACCAAGACGCTCTCCAAAGAGAACCAAAAGCCCAACAGCAACAGATAGT	360
Db	123	SerSerArgPro-----GlnAspArg	129
QY	361	CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCCTCCACCGGTGTGCT	420
Db	130	HisGlnLysIleTyrAsnSerArgGluGlyAspLeuIleAlaValProThrGlyValAla	149
QY	421	TTCTGCTCTACACGACCCACGACCTGATGTGTGTGTGTCTCTTCTTACTGACACCAAC	480
Db	150	TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn	169
QY	481	AACCAACGCAACACGCTTCGATCAGTTCCCGCAGAGATTCAATTTGGCTGGGAACACGGAG	540
Db	170	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	189
QY	541	CAAGAGTCTTTAAGTACCAGCAACAAAGCAGACAAAGCAGCAAGAAAGCTTACCATAT	600
Db	190	GlnGluPheLeuLysTyrGlnGlnGln-----	199
QY	601	AGCCCATACAGCCCGCAAGTCTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGCA	660
Db	200	-----GlyGly	201
QY	661	CAGCACAGCCGCGAGACGACGACGACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	720
Db	202	HisGlnSerGlnLysGlyLysHisGlnGlnGlnGluGluAsnGluGlySerIleLeu	221
QY	721	AGCGGCTTCACCGCGGAGTCTCTGGAACAAGCGCTTCCAGGTTGACGACAGACAGATAGTG	780
Db	222	SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla	240
QY	781	CAAAACCTAAGAGCGAGACCGAGAGTGAAGAGAGGAGGACCATTTGTGACAGTGGGGA	840
Db	241	LysAsnLeuGlnGlyLysGlnGluGlyGluAspLysGlyAlaIleValThrValLysGly	260
QY	841	GGCCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTGCCCGACGAAGAGAG	891
Db	261	GlyLeuSerValIleLysProThrAspGluGlnGlnArgProGlnGluGlu	280
QY	892	GAATAGTAGAAGATGATATGAATAC-----GATCAAGAG---GATAGAAAGG	936
Db	281	GluGluGluGluAspLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro	300
QY	937	CGTGGCAGGGGAGCAGAGCGAGGGGATGTTATTGAAGACGATCTGCACCGCAACT	996
Db	301	ArgGlySerGlnSerLysSerArgAsnGlyIleAspGluThrIleCysThrMetArg	320
QY	997	GCTAAAAAGACATTTGGTAGAAACAGATCCCTCGACATCTACAACCCCTCAAGCTGGTTCA	1056
Db	321	LeuArgHisAsnIleGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer	340
QY	1057	CTCAAACTGCCACGATCTCAACCTCTTAATCTTAGTGGCTTGGAGCTAGTGTCTGAA	1116
Db	341	ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGly	360
QY	1117	TATGGAATCTCTACAGGAATCATTTGTGTGCTCTACTACACACCAACACGACAGC	1176
Db	361	PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer	380

QY 1378 GAAAACTCCGTCATAGATAACCTCGCGAGGAGGTGGTTCARATTCATATGGCTCCAA 1437
 Db
 QY 473 ThSerSerValIleAsnAsnLeuProLeuAspValValAlaAlaThrPheAsnLeuGln 492
 Db
 QY 1438 AGGAGCAGCAAGCGAGCTTAAAGAACACAAACCCCTTCAAGTTCTTCGTTCCACCGTCT 1497
 Db
 QY 493 ArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProAlaArg 512
 QY 1498 CAGCAGTCTCCGAGGCTGTGGCT 1521
 Db
 QY 513 GlnSerGluAsnArgAlaSerAla 520
 Db
 RESULT 8
 FWPMLA
 legumin A precursor - garden pea
 C:Species: Pisum sativum (garden pea)
 C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 09-Jul-2004
 C:Accession: A22866
 R:Lycett, G.W.; Crox, R.R.D.; Shirsat, A.H.; Boulter, D.
 Nucleic Acids Res. 12, 4493-4506, 1984
 A:Title: The complete nucleotide sequence of a legumin gene from pea (Pisum sativum L.).
 A:Reference number: A22866; MUID:84247316; PMID:6330672
 A:Molecule type: DNA
 A:Residues: 1-517 <LYC>
 A:Cross-references: UNIPROT:P02857
 A:Experimental source: cv. Feltham First
 C:Comment: This protein, found in the seeds of many leguminous and nonleguminous plants,
 C:Genetics:
 A:Gene: leg A
 A:Introns: 96/1; 179/3; 388/3
 C:Superfamily: glycinin
 C:Keywords: seed; storage protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-332/Product: legumin A, alpha chain #status predicted <ALP>
 F:333-517/Product: legumin A, beta chain #status predicted <BET>

Alignment Scores:
 Pred. No.: 4.2e-99 Length: 517
 Score: 1440.00 Matches: 290
 Percent Similarity: 65.38% Conservatives: 67
 Best Local Similarity: 53.11% Mismatches: 99
 Query Match: 53.51% Indels: 90
 DB: 1 Gaps: 8

US-10-728-323-3 (1-1524) x FWPMLA (1-517)

QY 1 CGGCAGCAACCGGAGAGACGGCTGCCAGTTCAGCGCTCAATGCGCAGACCTGAC 60
 Db
 QY 23 ArgGluGlnProGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuProAsp 42
 Db
 QY 61 AATCGCATTGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACCAACGAGGATTC 120
 Db
 QY 43 AsnArgileGluSerGluGlyGlyLeulleGluThrTrpAsnProAsnAsnLysGlnPhe 62
 Db
 QY 121 GAATGCGCGCGCTCGCCCTCTCTCGTTAGTCTCGCGCGCAACGCCCTTCGTAGGCT 180
 Db
 QY 63 ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgPro 82
 Db
 QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGGATCTTTGGGTG 240
 Db
 QY 83 TyrTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyAsnGlyTyrPheGlyMec 102
 Db
 QY 241 ATATTCCTCGTGTCTTACACTATGAGAGCCTTCACACACAGGTCGTTCGATCTCAG 300
 Db
 QY 103 ValPheProGlyCysProGluThrPheGluGluPro-----GlnGlu 116
 Db
 QY 301 TCCCAAGACCCACAGAGCTCTCCAGAGGAGAGACCAAGCCCAACAGCAAGATAGT 360
 Db
 QY 117 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspArg 128
 Db
 QY 361 CACCAGAAGTGACCGTTTCGATAGGGGTGATCTCATTTGACGTTCACCGCGTGTGCT 420
 Db
 QY 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148
 QY 421 TTCGTGCTCTACCAACGACACGACGACGATGATGTTGCTGTTCTTCTACTGACACCAAC 480
 Db
 QY 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg 168
 QY 481 AACAAACGACCAACCGCTTCATCAGTTCCCGCAGGAGATTCATTTGCTGGGACACGGAG 540
 Db
 QY 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu 188
 QY 541 CAAGAGTCTTAAAGGTACCAGCAACAAAGCAGACAAAGCAGACGACGAAAGAGCTTACCATAT 600
 Db
 QY 189 GlnGluPheLeuGlnTyrGlnHisGln----- 197
 QY 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
 Db
 QY 661 CAGCAGCGCGCAGAGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 720
 Db
 QY 198 -----GlnGlyGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 212
 QY 721 AGCGCTTTCACGCGCGGAGTTCTCTGGAACAAAGCCCTTCAGGTTGACGACGACGACGATAGTG 780
 Db
 QY 213 SerGlyPheLysArgAspTyrLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 231
 QY 781 CAAAACCTTAAGAGCGCAGACCCGAGAGTGAAGAGAGGAGGCCCTTGTGACGAGGAGGA 840
 Db
 QY 232 AspArgLeuGlnGlyArgAsnGluAspGluGluLysGlyAlaIleValLysValLysGly 251
 QY 841 GCCTCAGAAATCTTGAGCCCGACAGATAGAAAG----- 870
 Db
 QY 252 GlyLeuSerIleIleSerProGluLysGlnAlaArgHisGlnArgGlySerArgGln 271
 QY 871 -----AGCGTCCCGCAGCAAGAGAG-----AGCGTCCCGCAGCAAGAGAGAG 891
 Db
 QY 272 GluGluAspGluAspGluLysGlnProArgHisGlnArgGlySerArgGlnGlu 291
 QY 892 GAATGATGATGAAGAGACGATCTGCGCAAGTGTATAAAGAGAAACATTCGTAGAACAGA 1023
 Db
 QY 332 AsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlyProSerSer 351
 QY 1024 TCCCTGACATCTACAAACCTCAAGCTGTTCACTCAAAACGACGACGACGACGACGACGACG 1083
 Db
 QY 352 SerProAspIleTyrAsnProGluAlaGlyArgIleLysThrValThrSerLeuAspLeu 371
 QY 1084 CTAACTACTAGTGGCTGGACCTAGTGTGATATGGAATCTCTACAGGATGATTC 1143
 Db
 QY 372 ProValLeuArgTyrLeuLysLeuSerAlaGluHisGlySerLeuHisLysAsnAlaMet 391
 QY 1144 TTTGTGCTCTACTACAAACCAACGACGACGACGACGACGACGACGACGACGACGACGACG 1203
 Db
 QY 392 PheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeuLysGlyArgAla 411
 QY 1204 CAGGTGCAAGTCTGGACAGCAACCGGCAACAGAGTGTACGACGAGGAGGCTTCAAGAGGT 1263
 Db
 QY 412 ArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGlyLeuGluAlaGly 431
 QY 1264 CAGGTGCTTGTGCTGCCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCAGAGATTC 1323
 Db
 QY 432 ArgAlaLeuThrValProGlnAsnTyrAlaValAlaAlaLysSerLeuSerAspArgPhe 451
 QY 1324 GAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCCAACCTCCCGGTGAAAAA 1383
 Db
 QY 452 SerTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSer 471
 Db

RESULT 10

S11003
glycinin G3 precursor - soybean
C;Species: Glycine max (soybean)
C;Date: 21-Nov-1993 #sequence_revision 19-Jan-1996 #text_change 03-May-1996
C;Accession: S11003
R;Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallan, B.J.; Fischer, R.L.;
Plant Cell 1, 313-326, 1989
A;Title: Characterization of the glycinin gene family in soybean.
A;Reference number: S10851; MUID:92393391; PMID:2485233
A;Accession: S11003
A;Molecule type: DNA
A;Residues: 1-484 <NIE>
A;Experimental source: variety Dare
C;Genetics:
C;Gene: GY3
C;Superfamily: glycinin
C;Keywords: storage protein
E;1-19/Domain: signal sequence #status predicted <SIG>
F;20-484/Product: glycinin G3 #status predicted <MAR>

Alignment Scores:
Pred. No.: 3.25e-98 Length: 484
Score: 1428.00 Matches: 290
Percent Similarity: 70.33% Conservative: 68
Best Local Similarity: 56.97% Mismatches: 95
Query Match: 53.07% Indels: 56
DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x S11003 (1-484)

```
QY 1 CGGCAGCAACCGGAGAGAACGCTGTCAGGTCCTCAGCGCTCAATGCGCAGACCTGAC 60
DB 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTCGAATCAGAGGGGGTTCATATTGACATCTGGACCCCAACACCGAGGATTC 120
DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGTCGCGCTCTCTCGCTTAGTCTCCGCGCGCAACGCGCTTCGTAGGCCT 180
DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATCTTTGGGTTG 240
DB 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCTCGTGTCTAGACACTATGAAGGCTTCACACACAAGTCTGTCGATCTCAG 300
DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACACCAAGACGCTCTCCAAGGAGAGAACCAAGCCAAACAGCAACAGAGATG 360
DB 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAGAGGTGCACCGTTTCGATAGGGTGATCTCATTCAGTTCACCGGTGTTGCT 420
DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCGGCTCTACAGCACCAGCAGACTGATGTTGTTGCTGTTCTTCTACTGACCAAC 480
DB 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACACGACACCAACGCTTGATCACTTCCCGAGGAGATTCATTTGGCTGGGAACACCGAG 540
DB 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTTCCTTAAGGTACCGACCAACAAAGCAGACAAAGCAGACGAGCAAGCTTACCATAT 600
DB 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCCGCAAGTACGCTTAGACAGAGAGCGTGATTTAGCCCTCGAGGA 660
```

```
DB 200 -----Gly 200
QY 661 CAGCACAGCCGACAGAACGAGCAGGACCAAGAGAGAAAGAAAGAGTGGAAACATCTTC 720
DB 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluAsnGluGlyGlySerIleLeu 220
QY 721 AGCGGCTTCAGCGCGGAGTCTCCGGAACAGCCCTCCAGGTTCCAGCACACAGATAGTG 780
DB 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY 781 CAAACCTAAGAGCGGAGACCGGAGAGTGAAGAGAGAGAGCCCATTTGTGACAGTGAAGGGA 840
DB 240 ArgLysLeuGlnGlnGluAsnGluGluGlyLysGlyAlaIleValThrValLysGly 259
QY 841 GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGAGCTGCCGACGAGAGAG 891
DB 260 GlyLeuSerValIleSerProThrGluGlnGlnArgProGluGluGluGlu 279
QY 892 GAATACATGACATGAATATGATACGATGAGAGAGATAGAAGCGTGCAGGGGAGGC 951
DB 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGGCGAGGGGAAATGCTATTGAAGAGAGCATCTGCACGCAAGTGTCTAAAGAAACATT 1011
DB 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAAACTGCCAAC 1071
DB 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
QY 1072 GATCTCAACCTTCTAATATTAGTGGCTTGGACCTAGTCTGAATATGGAATCTCTAC 1131
DB 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGGAATGCAATTGTTCTGCTCACTACAACCAACGACGACAGAGTGTACGAGGAG 1191
DB 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGAGGAG 1251
DB 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGTCACGTGCTTGTGGTCCACAGAACTTCGCCCTCGCTGGAAGTCCAG 1311
DB 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCGAAGCTTCCAATACGTGGCATTCAAGACAGACTCAAGGCCCGCAGCATAGCAACTC 1371
DB 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCCGGTGAATACTCCCTCATAGATAACCTGCCGAGGAGTGTGTGCAATATTCATATGGC 1431
DB 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCCTCAAGGAGGAGGAGGAGGAGGAGTGAAGAACACCAACCCCTTCAAGTCTTCGTT--- 1488
DB 452 LeuArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsn 471
QY 1489 CCACCTCTCAGCAG---TCTCCGAGG 1512
DB 472 ProGluSerGlnGlnGlySerProArg 480
```

RESULT 11

S14393
legumin A2 precursor - fava bean
C;Species: Vicia faba (fava bean)
C;Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S14393
R;Schlesier, B.; Bassuener, R.; van Hai, N.; Muentz, K.
Nucleic Acids Res. 18, 7146, 1990
A;Title: The cDNA derived primary structure of two distinct legumin A subunit precursors
A;Reference number: S14392; MUID:91088307; PMID:2263481

A;Accession: S14393
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-500 <SCH>
A;Cross-references: UNIPROT:Q99304; EMBL:X55014; NID:g22008
C;Superfamily: glycinin

Alignment Scores:

Pred. No.: 3 29e-97 Length: 500
Score: 1414.50 Matches: 280
Percent Similarity: 67.30% Conservative: 76
Best Local Similarity: 52.93% Mismatches: 100
Query Match: 52.56% Indels: 73
DB: 2 Gaps: 8

US-10-728-323-3 (1-1524) x S14393 (1-500)

QY	1	CGCAGAACCGAGGAGAACGCGTCCAGTTCACGGCTCAATCGCGAGAGACCTGAC	60
Db	23	ArgGluGlnSerGlnGlnAenGluCysGlnLeuGluArgLeuAspAlaLeuGluProAasp	42
QY	61	AATCGCATTGAATCAGAGGCGGTACATTGACACTTGGAAACCCCAACACAGGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrIrpAsnProAsnAsnArgGlnPhe	62
QY	121	GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT	180
Db	63	ArgCysAlaSerValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgAcpPro	82
QY	181	TTCTACTCCATGCTCCAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTG	240
Db	83	TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet	102
QY	241	ATATTCCCTGGTGTCTAGACACTATGAAGCCCTCACACAAAGTCTCGATCTCAG	300
Db	103	ValPheProSerCysProGluThrPheGluGluPro-----GlnGln	116
QY	301	TCCCAAGACCAACGAGCTCTCCAGGAGGAAGACCAAGGCCAACAGCAACGAGATGT	360
Db	117	SerGlu-----GlnGlyGlu-----GlyGlyArgTyrArgAaspSer	128
QY	361	CACCAGAGTGACCGTTTCGATAGGGTGATCTCATTCGAGTTCCTCCACCGGTGTGCT	420
Db	129	HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal	148
QY	421	TTCTGGCTCTACAAACACACACACTGATGTTGTGCTGTTCTCTTACTGACACCAAC	480
Db	149	PheIrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspIleGly	168
QY	481	AACAACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCAATTGCTGGGAACACGGAG	540
Db	169	SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	188
QY	541	CAAGAGTCTTAAGGTACCGAACCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT	600
Db	189	GlnGluPheLeuArgTyrGlnHisGln-----	197
QY	601	AGCCCATACGCCGCAAGTCCAGCTAGACAGAGAGCGTGAATTTAGCCCTCCAGGA	660
Db	197	-----	197
QY	661	CAGCAGACCGCAGAGAAACGAGCAGGACCAAGAAAGAAACGAAGGTGGAAACATCTTC	720
Db	198	-----GlnGlyGlyGlyGluGluGlnAspAsnAspGlyAsnAsnIlePhe	212
QY	721	AGCGGCTTACCGCGGAGTTCCTGGAAACAAAGCCCTCCAGCTTGACACAGACAGATAGTG	780
Db	213	SerGlyPheIlysArgAspPheLeuGluAspAlaLeuAsnVal---AsnArgHisIleVal	231
QY	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGCGCACTTGACAGTGGAGGA	840
Db	232	AspArgLeuGlnGlyArgAsnGluAspGluGlyGlyAlaIleValIlyValIlyGly	251

QY	841	GGCTTCAGAACTCTTGAGCCAGCATAGAAAG-----AGACGTGCCGACGAA	885
Db	252	GlyLeuSerIleIleThrProGluArgGlnAlaArgHisProArgGlySerArgGln	271
QY	886	GAAAGGAATACGATCAGATCAATATGAA-----	915
Db	272	GluGluAspGluAspGluAspGluGlyGluArgGlnProSerHisIlySerArg	291
QY	916	-----TACGATGAAGAGGATAGA-----AGCGTGGCAGGGAAGCAGA	954
Db	292	ArgGlyGluAspGluAspAspIlysgLulysArgHisSerGlnIlysgLysGluSerArgArg	311
QY	955	GGCAGGGGAATGGTATTGAAGAGACGATCTCCACCGCAAGTGCTAAAGAACAATTGGT	1014
Db	312	HisGlyAspAsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGly	331
QY	1015	AGAAACAGATCCCTGACATCTACACCTCAAGCTGGTTCACCTCAAACTGCCAAGAT	1074
Db	332	SerSerSerSerProAspIleTyrAsnProGlnAlaGlyArgIleIlyThrValThrSer	351
QY	1075	CTCAACTTCTAATCTTAGTGGCTTGACCTAGTCTGTAATATGAAATCTCTACAGG	1134
Db	352	LeuAspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLys	371
QY	1135	AATGCATTGTTGTCTCCTCACTACACCAACCAACGACACAGCATCATATATCGATTGAGG	1194
Db	372	AsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleLeuTyrAlaLeuLys	391
QY	1195	GGACGGCTCAGCTCGAAGTCTGTGACACAGCAAGCGCACAGATGTACACGAGGAGCTT	1254
Db	392	GlyArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGluGluLeu	411
QY	1255	CAAGAGGGTCAGCTGCTGTGTGTCACAGAACTTCGCCGTGCTGGAAAGTCCACAGAGC	1314
Db	412	GluAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSer	431
QY	1315	GAGAACTTCGAATACGTGCATTCAAGACAGACTCAAGCCCGACGATACCCAACTCGCC	1374
Db	432	AspArgPheThrTyrValAlaPheIlyThrAsnAspArgAlaGlyIleAlaArgLeuAla	451
QY	1375	GGTGAACCTCGTTCATAGATAACCTCGCGGAGGAGTGGTTCGCAATTCATATGGCTTC	1434
Db	452	GlyThrSerSerValIleAsnAspMetProValAspValValAlaAlaThrPheAsnLeu	471
QY	1435	CAAGGGAGCAGCGAAGGAGCTTAAGAACAAACCAACCCCTTCAAGTTCCTCGTTCACCG	1494
Db	472	GluArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProPro	491
QY	1495	TCTCAGCAGTCTCCGAGGCTGTGCT	1521
Db	492	ArgGluSerGlnLysArgAlaSerAla	500

RESULT 12

T06452

probable legumin A precursor - garden pea
C:Species: Pisum sativum (garden pea)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000

C:Accession: T06452

R:Lycett, G.W.; Croy, R.R.D.; Shirsat, A.H.; Richards, D.M.; Boulter, D.

Nucleic Acids Res. 13, 6733-6743, 1985

A:Title: The 5'-flanking regions of three legumin genes: comparison of the DNA sequences.

A:Reference number: Z15687; MUID:86041868; PMID:2997721

A:Accession: T06452

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-507 <LYC>

A:Cross-references: EMBL:X02982; NID:g20777; PIDN:CAA26720.1; PID:g4379378

C:Genetics:

A:Introns: 96/1; 179/3; 388/3

C:Superfamily: glycinin

C:Keywords: seed; storage protein

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-332/Product: legumin A, alpha chain #status predicted <ALP>

Alignment Scores:
 Pred. No.: 6.1e-95 Length: 507
 Score: 1384.00 Matches: 283
 Percent Similarity: 63.92% Conservative: 66
 Best Local Similarity: 51.83% Mismatches: 97
 Query Match: 51.43% Indels: 100
 DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x T06452 (1-507)

```

QY 1 CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCAGCGCCCTCAATGGCAGACCTGAC 60
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
QY 61 AATCGCATTAATCAGAGCGGTTACATTGAGACTTTGGAAACCCCAACACCGAGGATTC 120
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 43 AsnArgileGluSerGluGlyGlyLeuileGluThrTrpAsnProAsnAsnLysGlnPhe 62
QY 121 GAATGGCCGGCTCGCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCCT 180
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAAAGGGGATACTTTGGTTG 240
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 TyrTyrSerAsnAlaProGlnGluPheileGlnGlnGlyAsnGlyTyrPheGlyMet 102
QY 241 ATATTCCCTGGTTGCTTAGACTATGAAGAGCTTCACACAAAGTGTGCGATCTCAG 300
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 ValPheProGlyCysProGluThrPheGluPro-----GlnGlu 116
QY 301 TCCCAAGACCCACAGAGCTCTCCAGGAGAGAACCAAGCCACAGCAACAGATAGT 360
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 SerGlu-----GlnGlyGlu-----GlyArgArgTyrArgAspArg 128
QY 361 CACGAGAGGTGCACCGTTCCGATGAGGTGATCTCATTCAGTTCACCGGTGTTGCT 420
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148
QY 421 TTCTGGCTCTACACGACACGACACTGATGTGTTGCTGTTCTCTTACTGACACCAAC 480
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg 168
QY 481 AACACGACACACGCTTGATAGTTCCTCCAGAGATTCATTTGGCTGGGAACACGGAG 540
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu 188
QY 541 CAAGAGTTCTTAAGGTACACGACCAACAAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT 600
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 GlnGluPheLeuGlnTyrGlnHisGln----- 197
QY 601 AGCCCATACAGCCCGCAAGTGCAGCTTAGACAAGAGCGTGGAATTTAGCCCTCGAGGA 660
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 197 ----- 197
QY 661 CAGCAGACCGCGAGAGAACGAGCAGGACAGACAGAAAGAAAGAGTGGAAACATCTTC 720
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 -----GlnGlyGlyLysGlnGluGlnGluAsnGluGlyAsnAsnIlePhe 212
QY 721 AGCGGCTTACCGCGGAGTTCTCGGACAGACCTTCAGGTTCCAGCAGACAGATAGTG 780
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 SerGlyPheLysArgAspTyrLeuGluAspAlaPheAsnVal---AsnArgHisIleVal 231
QY 781 CAAAACTTAAGCGCAGACCGGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGGAGGGA 840
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGluGlyAlaIleValLysValLysGly 251
QY 841 GGCTCTAGATCTTGAGCCCGCAGATAGAAG----- 870
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 GlyLeuSerIleIleSerProGluLysGlnAlaArgHisGlnArgGlySerArgGln 271
QY 871 -----AGACGTCCCGAGGAAGAG 891
  |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Alignment Scores:
 Pred. No.: 2.84e-94 Length: 497
 Score: 1375.00 Matches: 276
 Percent Similarity: 68.80% Conservative: 79
 Best Local Similarity: 53.49% Mismatches: 107
 Query Match: 51.10% Indels: 54

RESULT 13

S14392

legumin A1 precursor - fava bean (fragment)

C:Species: Vicia faba (fava bean)

C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: S14392

R:Schlester, B.; Baasunen, R.; van Hai, N.; Muentz, K.

Nucleic Acids Res. 18, 7146, 1990

A:Title: The cDNA derived primary structure of two distinct legumin A subunit precursors

A:Reference number: S14392; MUID:91088307; PMID:2263481

A:Accession: S14392

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-497 <SCH>

A:Cross-references: UNIPROT:Q03971; EMBL:X55013; NID:g22006; PIDN:CAA38757.1; PID:g388251

C:Superfamily: glycinin

DB:	2	Gaps:	11
US-10-728-323-3 (1-1524) x S14392 (1-497)			
QY	25	TGCAGTTCAGGCGCTCAATGCGCAGAGACCTGACATCGCATTTGAATCAGAGGGCGGT	84
Db	1	CysGlnLeuGluArgLeuAspAlaLeuGluProAspAsnArgIleGluSerGluGly	20
QY	85	TACATTGAGACTTGGAAACCCCAACAACAGGAGTTGGAATGCGCCGCGTCCCTCTCT	144
Db	21	LeuIleGluThrIrpAsnProAsnAsnArgGlnPheArgCysAlaArgValAlaLeuSer	40
QY	145	CGTTAGTCTCCGCGCAACCGCTTCTGAGGCTTTCTACTCCAATGCTCCCCAGGAG	204
Db	41	ArgAlaThrLeuGlnArgAsnAlaLeuArgArgProIlyrSerAsnAlaProGlnGlu	60
QY	205	ATCTTATCCAGCAGAGGAGATACTTTGGTGTGATATTCCTCGTGTCTCTAGACAC	264
Db	61	IleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMetValPheProSerCysProGluThr	80
QY	265	TATGAAGAGCTCACACACAAGTTCGTCGATCTCAGTCCCAAGACACCAAGACGCTC	324
Db	81	PheGlnGluPro-----GlnGlnSerGlu-----	88
QY	325	CAAGGAGAACCAAAACCAACGACGATAGTACACAGAGGTGCACCGCTTCGAT	384
Db	89	GlnGlyGlu-----GlyArgArgTyrArgAspSerHisGlnLysValAsnArgPheArg	106
QY	385	GAGGTGATCTCAATGAGTTCACCGGTGTCTTCTTCTGCTCTACACGACCAACGAC	444
Db	107	GlnGlyAspIleAlaValProThrGlyIleValPheTrpMetTyrAsnAspGlnAsp	126
QY	445	ACTGATCTTCTGCTGTTCTTCTACTGACACCAACAACACACACAGCTTGATCAG	504
Db	127	IleProValIleAlaIleSerLeuThrAspThrGlySerSerAsnGlnLeuAspGln	146
QY	505	TTCCCGAGGAGATCAATTTGGCTGGGAAACGCGAGCAAGAGTCTTAAAGGTACCAACAA	564
Db	147	MetProArgArgPheTyrLeuAlaGlyAsnGlnGluGlnPheLeuArgTyrGlnHis	166
QY	565	CAAGCGACAAAGCAGACGAGAGCTTACCATATAGCCCATACAGCCCGGAAAGTCAG	624
Db	167	GlnGlnGlyValLysGluGlu-----GlnAspAsn	176
QY	625	CCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACGACACAGCGCGCAGAGAACGACGA	684
Db	177	AspGlyAsnGlnGlnGluPheLeu---ArgTyrGlnHis-----ArgGlnGlyVal	193
QY	685	GGACAAAGAAGAAACGAAGTGGAAACATCTTCAGCGGCTTCACGCGGAGTTCTCTG	744
Db	194	LysGluGlnAspAsnAspGlyAsnAsnIlePheSerGlyPheAsnArgAspPheLeu	213
QY	745	GAACAGCTTCAGGTTGACGACGACAGATAGTGCMAACCTAAGAGCGCGAGCCGAG	804
Db	214	GluAspAlaPheAsnVal---AsnArgHisIleValAspArgLeuGlnGly-----Arg	230
QY	805	AGTGAGAAGAGGAGCATTGTGACAGTGCAGGCGGCTCAGATCTTGAGCCCCAGAT	864
Db	231	AsnGlnGluArgGlyAlaIleValLysValLysGlyGlyLeuSerIleIleThrProPro	250
QY	865	AGAAAGAGA-----CGTCCGACGAGAAAGAGGAGGAATACGATGAA	903
Db	251	GluArgGlnAlaArgHisProArgGlySerArgGlnGluGluAspGluAspGlu	270
QY	904	GATGAATATGAA-----TACGATGAAGAG	927
Db	271	AspGluLysGluGluArgGlnProSerHisHisLysSerArgArgGlyGluAspGluAsp	290
QY	928	GATAGA-----AGCGGTGGCAGGGAAGCAGCGGAGGGGAATGTATT	972
Db	291	AspLysGluLysArgHisSerGlnLysGlyLeuSerArgArgHisGlyAspAsnGlyLeu	310
QY	973	GAAGAGCGATCTGCACCGCGCAAGTCTGCTAAAGAAACATTGGTAGAAACAGATCCCTGAC	1032

Db	311	GluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlySerSerSerProAsp	330
QY	1033	ATCTACACCCCTCAGCTGGTTCACCTCAAACTGCCAAGCATCTCAACCTTCTAATACTT	1092
Db	331	IleTyrAsnProGlnAlaGlyArgIleIleYsThrValThrSerValAspLeuProValLeu	350
QY	1093	AGGTGGCTTGGACCTAGTGTGAATATGAAATCTCTACAGGAATGCATTTGTTGTCGCT	1152
Db	351	ArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLysAsnAlaMetPheValPro	370
QY	1153	CACTCAACACCAACGACACACATCATATATCGATTGAGGGAGCGGCTCACGTGCAA	1212
Db	371	HisTyrAsnLeuAsnAlaAsnSerValLeuTyrAlaLeuLysGlyArgAlaArgLeuGln	390
QY	1213	GTCTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCACAGAGGCTCACGTGCTT	1272
Db	391	ValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaGlyArgAlaLeu	410
QY	1273	GTGGTGCCACAGAACTTCGCGCTGGAAAGTCCAGAGCGAGAACTTCGAATACGTG	1332
Db	411	ThrValProGlnAsnTyrValValAlaLysSerLeuSerAspArgPheThrTyrVal	430
QY	1333	GCATTCAAGACAGACTCAAGCCCATAGCAACCTCGCCGTCGAAACTCCGTCATA	1392
Db	431	AlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSerSerValIle	450
QY	1393	GATAACCTCCCGAGGAGGTGCTTGCATTCATATGCTCCAAAGGAGCAGCGCAAGG	1452
Db	451	AsnAspLeuProLeuAspValValAlaAlaThrPheAsnLeuGluArgAsnGluAlaArg	470
QY	1453	CAGCTTAAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAG	1500
Db	471	GlnLeuLysPheAsnAsnProSerArgPheLeuValProProArgGlu	486
RESULT 14			
S51941			
N: prunin 1 precursor - almond			
C: Species: Prunus dulcis (almond)			
C: Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 09-Jul-2004			
C: Accession: S51941; S42474			
R: Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.			
Plant Mol. Biol. 27, 205-210, 1995			
A: Title: Molecular characterization of cDNAs corresponding to genes expressed during alme			
A: Reference number: S51940; MUID: 95170003; PMID: 7865791			
A: Accession: S51941			
A: Status: nucleic acid sequence not shown			
A: Molecule type: mRNA			
A: Residues: 1-551 <GAR>			
A: Cross-references: UNIPROT: Q43607; EMBL: X78119			
A: Note: the source is designated as Prunus amygdalus			
R: Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.			
submitted to the EMBL Data Library, March 1994			
A: Description: Molecular characterization of cDNAs corresponding to proteins related to			
A: Reference number: S42473			
A: Accession: S42474			
A: Molecule type: mRNA			
A: Residues: 1-60, 'G', 62-551 <GAW>			
A: Cross-references: EMBL: X78119; NID: g460805; PIDN: CAA55009.1; PID: g460806			
A: Note: the source is designated as Prunus amygdalus			
C: Superfamily: glycylin			
C: Keywords: seed; storage protein			
F: 1-20/Domains: signal sequence #status predicted <SIG>			
F: 21-367/Product: prunin 1 alpha chain #status predicted <MAT1>			
F: 368-551/Product: prunin 1 beta chain #status predicted <MAT2>			
F: 108-374/Disulfide bonds: #status predicted			
Alignment Scores:			
Pred. No.:	7, 47e-72	Length:	551
Score:	1073.50	Matches:	228
Percent Similarity:	57.22%	Conservative:	89
Best Local Similarity:	41.16%	Mismatches:	160

Query Match: 39.89% Indels: 77
DB: 2 Gaps: 11
US-10-728-323-3 (1-1524) x S51941 (1-551)

QY 4 CAGCAACCGGAGGAGACCGCTGCCAGTTCAGCGCCTCAATGCCAGAGACCTGACAAAT 63
DB 25 GlnLeuSerProGlnAenGlnCysGlnLeuAenGlnLeuGlnAlaAenGlnProAenAen 44
QY 64 CGCATTGAATCAGAGCGCGCTTACATTGAGACTTGGAAACCCCAACAAACAGAGAGTTCGAA 123
DB 45 ArgIleGlnAlaGluAlaGlyGlnIleGluThrTrpAenPheAenGlnGluAenPheGln 64
QY 124 TGGCCGCGCTCCCTCTCTCGCTTGTAGTCTCCGCCGCAACGCCCTTCGTAGGCTTTC 183
DB 65 CysAlaGlyValAlaAlaSerArgIleThrIleGlnArgAenGlyLeuHisLeuProSer 84
QY 184 TACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAGGAGGAGGATATCTTGGGTGATA 243
DB 85 TyrSerAenAlaProGlnLeuIleTyrIleValGlnGlyArgGlyValLeuGlyAlaVal 104
QY 244 TTCCTCGTGTCTAGACACTATGAAGAG----- 273
DB 105 PheSerGlyCysProGlnThrPheGluGluSerGlnGlnSerSerGlnGlnGlyArgGln 124
QY 274 ---CCTCACACAAAGTCTCGATCTCAGTCTCCAAAGACCAACCAAGACGCTCCAAAGGA 330
DB 125 GlnGluGlnGluGlnGluArgGlnGlnGlnGlnGlnGlyGluGlnGlyArgGlnGlnGly 144
QY 331 GAAGACCAAGCCCAACAGACGA----- 354
DB 145 GlnGlnGluGlnGlnGlnGluArgGlnGlyArgGlnGlnGlyArgGlnGlnGlnGluGlu 164
QY 354 ----- 354
DB 165 GlyArgGlnGlnGluGlnGlnGlnGlnGlnGlyArgProGlnGlnGlnGlnGlnPhe 184
QY 355 -----GATAGTCACCAAGAGTGCACCGCTTCGATGAGGTGATCTCAATGCGAGTT 405
DB 185 ArgGlnLeuAenAenGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 204
QY 406 CCACCGGTGCTTCTTCTGCTGTACACAGCAGCAGCAGCAGCAGTGTGTTGCTGTTCT 465
DB 205 ProAlaGlyValAlaTyrTrpSerTyrAenAenGlyArgGlnGlnGlnGlnGlnGlnGln 224
QY 466 CTCTACTGACACCAACCAACAGCAGCAGCTTGTATCATGCTCCCGAGAGATTCAAATTG 525
DB 225 LeuPheHisValSerSerAenHisAenGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 244
QY 526 GCTGGGAACACGAGCAAGAGTCTTTAAGGTACCAGCAACCAAGCAGACCAAGCAGACGA 585
DB 245 AlaGlyAenProGluAenGluPhe-----AenGlnGlnGly----- 256
QY 586 AGAAGCTTACCATTATAGCCCATACAGCCCGCAAGTCTAGCAGTACACAGAGAGCGGTGAA 645
DB 257 -----GlnSerGlnProArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 266
QY 646 TTTAGCCCTCGAGCAGCAGCAGCCGC-----AGAGAACGAGCAGGACAGAGAGAA 696
DB 267 GlyArgPro---GlyGlnHisGlnGlnProPheGlyArgProArgGln---GlnGluGln 284
QY 697 GAAACCAAGGTGGAAACATCTTACGCGCTTTCAGCGCGAGTCTTCGGAAACAGACCTTC 756
DB 285 GlnGlyAenGlyAenAenValPheSerGlyPheAenThrGlnLeuLeuAlaGlnAlaLeu 304
QY 757 CAGGTTGACGACAGACAGATAGTGCAAACTTAAGAGCGGAGACCGAGAGTGAAGAGAG 816
DB 305 AenValAenGlu---GluThrAlaAenAenLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 323
QY 817 GAGCCATTGTGACAGTGGAGGCGCTCAGAACTTTGAGCCCGCAGATAGAAAGAGCGT 876
DB 324 -----IleIleGlnValArgGlyAenLeuAenAenPheValGlnProProArgGlyArgGln 341

```

US-10-728-323-3 (1-1524) x S11004 (1-560)
QY 19 AAGCGTCCAGTCCAGCGCTCAATGCGCAGACCTGACATCGCATTTGAATCAGAG 78
Db 30 ASnGluCysGlnLeuAenAenLeuAenAlaLeuGluProAspHisArgValGluSerGlu 49
QY 79 GCGGTTTACATTGAGACTTGGAAACCCCAACAACAGGAGTTTCAATGCGCGCGTCGCC 138
Db 50 GlyGlyLeuIleGlnThrTrpAenSerGlnHisProGluLeuLysCysAlaGlyValThr 69
QY 139 CTCTCGCTTAGTCTCCGCCGCAACGCCCTTCGTAGGCTTTCTACTCCAATGCTCCC 198
Db 70 ValSerLysLeuThrLeuAenArgAenGlyLeuHisLeuProSerTyrSerProTyrPro 89
QY 199 CAGGAGATCTTCAATCCAGCAAGGAGGATATCTTTGGGTGATATTCCTCGGTCTCCT 258
Db 90 ArgMetIleIleAlaGlnGlyLysGlyAlaLeuGlyValAlaIleProGlyCysPro 109
QY 259 AGACACTATGAAGAGCTCACACACAAGTCTCGATCTCAGTCCCAAGACCACCAAGA 318
Db 110 GluThrPheGluGluProGlnGlnSerAsnArgArgGlySerArg 125
QY 319 CGTCTCCAGGAGAGAACCAACCAACAGCAACGATAGTCCAGCAAGGTGCACCGT 378
Db 126 -----SerGlnLysGlnGlnLeuLeuAspSerHisGlnLysIleArgHis 140
QY 379 TTCGATGAGGAGTCTCATTTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 498
Db 161 GlyAspGluProValValAlaIleSerLeuLeuAspThrSerAsnPheAenAsnGlnLeu 180
QY 499 GATCAGTTCGCCAGGAGATTCATTTGCTGGGAAACAGGAGCAAGAGTTCTTAAGGTAC 558
Db 181 AspGlnThrProArgValPheTyrLeuAlaGlyAsnProAspIleGluTyr 197
QY 559 CAGCAACAAGCAGACAAAGCAGCAGCAAGAGCTTACCATATAGCCCATACAGCCGCAA 618
Db 198 -----ProGluThrMetGln 202
QY 619 AGTCAGCTAGACAAGAAGCGTGAATTTAGCCCTCGAGCAGACACAGCCGAGAGAA 678
Db 203 GlnGlnGlnGlnLysSerHisGlyArgLysGlnGlyHisGln 219
QY 679 CGAGCAGGACAAGAAGAAAACGAGGTGGAAACATCTTCAGCGCTTTCAGCCGGAG 738
Db 220 -----GlnGluGluGluGluGluGlySerValLeuSerGlyPheSerLysHis 236
QY 739 TTCCTGGAACAAGCCTTCCAGGTTGACGACAGACAGATAGTGCAAAACCTAAGAGCGAG 798
Db 237 PheLeuAlaGlnSerPheAenThrAenGlu ---AspIleAlaGlnLysLeu 253
QY 799 ACCGAGAGTGAAGAAGGAGGAGCCATTGTGACAGTGGGGAGGCGCTCAGAATCTTCGAG 858
Db 254 SerProAspAspGluArgLysGlnIleValThrValGluGlyLeuSerValIleSer 273
QY 859 CCAGATAGAAAAGAGAGCTGCCGCAAGAAAGAGAAATACGATGAA 903
Db 274 ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu 293
QY 904 -----GATGAATAT 912
Db 294 GlnIleProSerHisProProArgArgProSerHisGlyLysArgGluGlnAspGluAsp 313
QY 913 GAATACGATGAAGAGGATAGA ---AGCGTGGCAGGAGGAGCAGCAGGC 957
Db 314 GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgGluGln 333
QY 957 ----- 957

```

```

Db 334 AspGlnAspGlnAspGluAspGluAspGluAspGlnProArgLysSerArgGlu 353
QY 958 -----AGGGGG 963
Db 354 TrpArgSerLysLysThrGlnProArgArgProArgGlnGluGluProArgGluArgGly 373
QY 964 -----AATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAGAACATT 1011
Db 374 CysGluThrArgAenGlyValGluGluAenIleCysThrLeuLysLeuHisGluAenIle 393
QY 1012 GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTGTCTCACTCAAACTGCCAAC 1071
Db 394 AlaArgProSerArgAlaAspPheTyrAsnProLysAlaGlyArgLysSerThrLeuAen 413
QY 1072 GATCTCAACCTTCTAATCTTAGTGGCTTGACCTAGTGTGCTGAATATGGAATCTCTAC 1131
Db 414 SerLeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnTyrValValLeuTyr 433
QY 1132 AGGAATGCAATTTGTTGCTCCTACTACACCAACGACACAGCATCATATATCGAATTG 1191
Db 434 LysAenGlyIleTyrSerProHisTrpAenLeuAenAlaAenSerValIleTyrValThr 453
QY 1192 AGGGACGGCTCAGTGCAGTCTGGACACGACGCAACAGAGTGTACGACGAGGAG 1251
Db 454 ArgGlyGlnGlyLysValArgValValAenCysGlnGlyAsnAlaValPheAspGlyGlu 473
QY 1252 CTTCAAGAGGGTCACTGCTGTGTGTGTCACAGAACTTCGCGTCTGCTGGAAGTCCCAG 1311
Db 474 LeuArgArgGlyGlnLeuLeuValValProGlnAenPheValValAlaGluGlnAlaGly 493
QY 1312 AGCGAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTTC 1371
Db 494 GluGlnGlyPheGluTyrIleValPheLysThrHisAenAlaValThrSerTyrLeu 513
QY 1372 GCGGTGAAAACTCCGTATAGATAACCTGCGGAGGAGTGTGTCGAATTCATATGCC 1431
Db 514 -----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyrAen 530
QY 1432 CTCCAAAGGGAGCAGGAGGAGCTTAAGAAACAACACCCCTCAAGTTCCTCGTTCCA 1491
Db 531 LeuArgArgGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro 550
QY 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 551 ProLysGluSerGlnArgArgValValAla 560

```

Search completed: August 24, 2005, 10:07:00
Job time : 42.0812 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 24, 2005, 03:57:26 ; Search time 93.2661 Seconds
(without alignments)
16735.100 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US10728323/runat 23082005 124355 29210/app_query.fasta 1.4757
-DB=Uniprot 03 -Qfmt=fastan -SUFFIX=n2p_rup -MINMATCH=0_1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pept -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10728323 @CGN 1.1 291 @runat 23082005 124355 29210 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPOP=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	99.3	507	2	O82580
2	2479	92.1	529	2	O9FZ11
3	2455	91.2	536	2	O647H4
4	2451.5	91.1	537	2	O647H3
5	2444	90.8	538	2	O8LKN1
6	2436	90.5	536	2	O6T2T4
7	2421	90.0	530	2	O9SQH7
8	1834	68.2	510	2	O61WG5
9	1651	61.4	484	2	O647H2
10	1502	55.8	481	1	GLC3_SOYBN
11	1499.5	55.7	485	1	GLC2_SOYBN
12	1496	55.6	495	1	GLC1_SOYBN
13	1489	55.3	481	2	O852U5
14	1488.5	55.3	482	2	O852U4
15	1457.5	54.2	498	2	O41702
16	1455	54.1	520	1	LEG2_PEA

17	1443	53.6	517	2	O9T0P5
18	1440	53.5	517	1	LEG2_PEA
19	1431.5	53.2	482	2	O41676
20	1414.5	52.6	500	2	O99304
21	1375	51.1	497	2	O03971
22	1186	44.1	496	2	O9SMJ4
23	1110.5	41.3	515	2	O8W1C2
24	1074.5	39.9	488	2	O41128
25	1069.5	39.7	551	2	O43607
26	1059	39.4	542	2	O8LK20
27	1049	39.0	560	2	O9S9D0
28	1034.5	38.4	457	2	O8GZP6
29	1032	38.4	503	1	LEGJ_PEA
30	999.5	37.1	566	2	O24234
31	998	37.1	485	2	O41703
32	996.5	37.0	484	1	LEG4_VICFA
33	986	36.6	517	2	O39922
34	986	36.6	517	2	O7GC77
35	985	36.6	517	2	O9SB12
36	981.5	36.5	564	2	O43673
37	980	36.4	517	2	P93707
38	980	36.4	517	2	P93708
39	978.5	36.4	219	2	O8L103
40	975	36.2	563	2	O9SB11
41	974	36.2	563	2	O39921
42	973.5	36.2	562	1	GLC4_SOYBN
43	968	36.0	504	2	O43608
44	951.5	35.4	662	2	O647H1
45	946.5	35.2	562	2	O43452

ALIGNMENTS

RESULT 1

O82580	ID	O82580	PRELIMINARY;	PRT;	507 AA.
AC	O82580;				
DT	01-NOV-1998	(TREMREL. 08, Created)			
DT	01-NOV-1998	(TREMREL. 08, Last sequence update)			
DT	01-MAR-2004	(TREMREL. 26, Last annotation update)			
DE	Glycinin (Fragment)				
GN	Name=Arah3;				
OS	Arachis hypogaea (Peanut)				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;				
OC	Arachis				
OX	NCBI_TaxID=3818;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99146968; PubMed=10021462;				
RA	Rabjohn P., Helm E.M., Stanley J.S., West C.M., Sampson H.A.,				
RA	Burks A.W., Bannan G.A.;				
RT	"Molecular cloning and epitope analysis of the peanut allergen Ara h				
RT	3.";				
RL	J. Clin. Invest. 103:535-542(1999).				
DR	EMBL; AF093541; AAC63045.1; -				
DR	HSSP; P04776; IUD1.				
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.				
DR	InterPro; IPR006045; Cupin.				
DR	InterPro; IPR007113; Cupin-region.				
DR	InterPro; IPR011051; RmlC-like_cupin.				
DR	InterPro; IPR006044; Seedstore_11s.				
DR	Pfam; PF00190; Cupin; 2.				
DR	PRINTS; PR00439; 11SGLOBULIN.				
DR	PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.				
FT	NON_TER 1				
SQ	SEQUENCE 507 AA; 56349 MW; F3FB38BC3CB82DED CRC64;				

Alignment Scores:

Pred. No.: 5,12e-179 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 99.33%		Indels: 0
DB: 2		Gaps: 0
US-10-728-323-3 (1-1524) x 082580 (1-507)		
QY	1	CGGACGACCGGAGGAGAACGGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC 60
DB	1	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProasp 20
QY	61	AATCGCATTAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACAACAGGAGTTC 120
DB	21	AsnArgileGluSerGluGlyGlyIleGluThrTrpAsnProAsnAenGlnGluPhe 40
QY	121	GAATGCGCGCGGTGCGCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAenAlaLeuArgGPro 60
QY	181	TTCTACTCCAATGCTCCCCAGGAGATCTTCATCCAGCAAGGAAGGGGATATTTGGGTTG 240
DB	61	PheTyrSerAenAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY	241	ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACAACAAGTGTGTCATCTCAG 300
DB	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 100
QY	301	TCCCAAGACCAACAAGACGTCTCCAAGGAGAGAACCAAGCCAAACAGCAACGAGATAGT 360
DB	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY	361	CACCAGAAGTGCACCGTTTCGATGAGGGTGTATCTCATTTGCAGTTCCTCCACCGGTGTGCT 420
DB	121	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY	421	TTCTGGCTCTACAACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACAAC 480
DB	141	PheTrpLeuTyrAenAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY	481	AACAACGACCAACGAGTTCATGATTCCTCCAGGAGATTCATTTGGCTGGGAACGCGAG 540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY	541	CAAGAGTCTTAAGTACCAGCAACAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
DB	181	GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200
QY	601	AGCCCATACAGCCGCAAGTTCAGCTAGACAGAGAGCGTCAATTTAGCCCTCCAGGA 660
DB	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY	661	CAGCAGACCGCAGAGAACGAGCAGGACAGACAGAAAGAAACCAAGGTGGAACATCTTC 720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluGluGluGlyAsnIlePhe 240
QY	721	AGCGGTTCACGCGGAGTTCCTGGAAACAAGCCTTCAGGTTCCACACAGACAGATAGTG 780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY	781	CAAAACCTAAGACGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTACGGGA 840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGlyAlaIleValThrValArgGly 280
QY	841	GGCTCTCAGAACTTTAGCCCCAGATAGAAAGAGACGTGCCGACCAAGAAAGAGGAATACGAT 900
DB	281	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 300
QY	901	GAAGATGAATATGAATACGATGAAGAGGATAGAAAGCGGTGGCAGGGGAACGAGACGAG 960
DB	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY	961	GGGAATGGTATTGAAGAGAGCATCTGCACCGCAAGTCTTAAAGAACATTGTGTAACAAC 1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340
QY	1021	AGATCCCTGCATCTACAAACCTCAAGCTGGTTCACCTCAAACCTGCCAACGATCTCAAC 1080
DB	341	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY	1081	CTTCTAATACTTAGTGGCTTGACCTAGTCTGTAATATGGAAATCTCTACAGGAATGCA 1140
DB	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380
QY	1141	TTGTTTGTTCGCTCAGTACCAACCAACGACACACGATCATATATCGATTGAGGGACGG 1200
DB	381	LeuPheValAlaHisTyrAsnThrAenAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY	1201	GCTCAGCTCCAAGTCTGTGACAGCAACGCAACAGAGTGTAGCAGCAGAGCTTCAAGAG 1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluGlnGlu 420
QY	1261	GGTACAGTCTGTTGGTGCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAAC 1320
DB	421	GlyHisValLeuValValProGlnAenPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY	1321	TTCCGATACGTGGCATTCAGACACACTCAAGGCCCGCATAGCCAACTCCCGGTGAA 1380
DB	441	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY	1381	AATCTCGTCATAGATAAACCCTGCGGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAG 1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY	1441	GAGCAGGCAAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAenProPheLysPhePheValProSerGln 500
QY	1501	CAGTCTCCGAGCGCTGTGGCT 1521
DB	501	GlnSerProArgAlaValAla 507
RESULT 2		
ID	Q9FZ11	PRELIMINARY; PRT; 529 AA.
AC	Q9FZ11	
DC	01-MAR-2001 (TrEMBLrel. 16, Created)	
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	Gly.	
GN	Name=Glyl;	
OS	Arachis hypogaea (Peanut).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;	
OC	Arachis.	
OX	NCBI_TaxID=3818;	
RN	[1]_TaxID=3818;	
RP	SEQUENCE FROM N.A.	
RA	Jain A.K., Basha S.M.;	
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF125192; AAG01363.1; -.	
DR	HSSP; P04776; IUCX.	
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.	
DR	InterPro; IPR006045; Cupin.	
DR	InterPro; IPR007113; Cupin region.	
DR	InterPro; IPR011051; RmC like cupin.	
DR	pfam; PF00190; Seedstore_1ls.	
DR	PRINTS; PR00439; 11SGLOBULIN.	
DR	PROSITE; PS00305; 11S SBED STORAGE; UNKNOWN_1.	
SQ	SEQUENCE 529 AA; 60448 MW; BB6F25BC1D6E06A1 CRC64;	
Alignment Scores:		
Pred. No.:	2,11e-165	Length: 529
Score:	2479.00	Matches: 475
Percent Similarity:	95.28%	Conservative: 9
Best Local Similarity:	93.50%	Mismatches: 20

Query Match:	92.12%	Indels:	4
DB:	2	Gaps:	2
US-10-728-323-3 (1-1524) x Q9F211 (1-529)			
QY	1	CGGCAGCAACCGGAGGAGAACGGCTGCGAGTTCCAGCGCCTCAATGCGCAGACCTGAC	60
DB	25	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnA-gProAep	44
QY	61	AATCGCATTAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACACAGAGGATTC	120
DB	45	AsnArgLeuGluSerGluGlyGlyTyriIeGluThrPheAenProAenAenGlnGluPhe	64
QY	121	GAATGCGCGCGCTGCGCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGSCCT	180
DB	65	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	84
QY	181	TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTTG	240
DB	85	PheTyrsrAenAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyriPheGlyLeu	104
QY	241	ATATTCCCTGTTGCTCTAGACACTATGAAGAGCTCACACAAAGTCTCGATCTCAG	300
DB	105	IlePheProGlyCysProSerThrTyriGluGluProAlaGlnGlnGlyA-gHisGln	124
QY	301	TCCCAAGACCAACAGAGCTCTCAAGGAGAGAACCAAGC---CAACAGCAACGAGAT	357
DB	125	SerGlnArgAlaProArg-gPheGluGlyCyluAspGlnSerGlnGlnGlnGlnAep	144
QY	358	AGTCACAGAGGTGCACCGTTTTCGATGAGGGTGATCTCATTCGAGTTCCACCGGTGT	417
DB	145	SerHisGlnIysValA-gA-gPheAspGluGlyAspLeuIleAlaValProThrGlyVal	164
QY	418	GCTTTCTGCTCTACACGACGACACTGATGTTGTTGTTGTTCTTCTTACTGACAC	477
DB	165	AlaLeuTrpMetPheAenAepHisAepThrAspValValAlaValSerLeuThrAepThr	184
QY	478	AACAAACACCAACACGCTTGATCAGTTCCCGAGGAGATTCATTTGGCTGGGAACAG	537
DB	185	AsnAenAenAepAenGlnLeuAepGlnPheProArgArgPheAenLeuAlaGlyAenHis	204
QY	538	GAGCAAGAGTCTTAAGGTACCGACCAACAGCAGACAAAGCAGACCAAGAGCTTACCA	597
DB	205	GluGlnGluPheLeuA-gTyriGlnGln-----GlnSerArgArgSerLeuPro	221
QY	598	TATAGCCCATACAGCCCGCAAGTGCAGCTAGACAAGAGCGGTGAATTTAGCCCTCGA	657
DB	222	TyriSerProTyriSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg	241
QY	658	GGACAGCACCGCCAGACAGCAGGACGAGCAAGAAAGAAACGAAGGTGGAAACATC	717
DB	242	GlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluAenGluGlyGlyAenIle	261
QY	718	TTACAGCGCTTCAGCGCGAGTTCTCTGAAACAAGCTTCAGGTTGACGACAGACATA	777
DB	262	PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAepA-gGlnIle	281
QY	778	GTGCAAAACCTTAAGAGCGGACCGCAGAGTCAAGAGAGGGAGCCATTGTGACAGTGAGG	837
DB	282	ValGlnAenLeuArgGlyGluAenGluSerGluGluGlyAlaIleValThrValLys	301
QY	838	GGAGGCTCAGATCTTTGAGCCAGATAGAAAGAGAGCTGCCGACGAAGAGGAATAC	897
DB	302	GlyGlyLeuArgIleLeuSerProAepArgIysArgGlyAlaAepGluGluGluGluTyri	321
QY	898	GATGAAGCATGAATGATACGATGAAGAGATAGAGGCTGCGAGGGAAGCAGAGGC	957
DB	322	AspGlnAepGluTyriGluTyriAepGluAepArgArgGlyArgGlySerArgGly	341
QY	958	AGGGGGAATGATTGAAGAGACCATCTGCACCCCAAGTGTCTAAAGAAACATTGGTAGA	1017
DB	342	ArgGlyAenGlyIleGluGluThrIleCysThrAlaSerValIysIysAenIleGlyArg	361
QY	1018	AACAGATCCCCTGACATCTACACCTCAAGCTGGTTCACCTCAAACTGCCACGATCTC	1077
DB	362	AsnArgSerProAepIleTyriAenProGlnAlaGlySerLeuIysThrAlaAenAepLeu	381
QY	1078	AACCTTTCTAATACTTAGGTGGCTTGACCTAGTGTGCTGAATATGAAATCTCTACAGAA	1137
DB	382	AsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyriGlyAenLeuTyriA-gAen	401
QY	1138	GCATTGTTTCTCCTCACTACAAACACCAACGACACAGCATATATATCATGATTGAGGGA	1197
DB	402	AlaLeuPheValProHisTyriAenThrAenAlaHisSerIleIleTyriAlaLeuArgGly	421
QY	1198	CGGGCTCAGTGCAGTCTGTCGACCAACGCGCAACAGAGTGACGACGAGGAGCTTCAA	1257
DB	422	ArgAlaHisValGlnValValAspSerAsnGlyAenArgValTyriAepGluGluLeuGln	441
QY	1258	GAGGTCACGCTGCTGTCGTCACAGAACTTCGCGCTCTCGTGAAGTCCACGAGCGAG	1317
DB	442	GluGlyHisValLeuValValProGlnAenPheAlaValAlaGlyIysSerGlnSerAep	461
QY	1318	AACCTTCGAATACGTGCATTCAAGACAGACTCAAGGCCACGATAGCCAACTCGCCGCT	1377
DB	462	AsnPheGluTyriValAlaPheIysThrAepSerArgProAenIleAlaAenPheAlaGly	481
QY	1378	GAATACTCCGTCATAGTAACTCCGCGAGGAGGTGTCGAATTCATATGCGCTCCAA	1437
DB	482	GluAenSerIleIleAepAenLeuProGluGluValValAalaAenSerTyriGlyLeuPro	501
QY	1438	AGGAGCAGGAGGAGGAGCTTAAAGAACAAACACCCCTTCAAGTCTCTGTTCCACCGCT	1497
DB	502	ArgGluGlnAlaArgGlnLeuIysAenAenAenProPheIysPheValProProSer	521
QY	1498	CAGCAGTCTCCGAGGCTGTGGCT 1521	
DB	522	GlnGlnSerLeuArgAlaValAla 529	
RESULT 3			
Q647H4	PRELIMINARY;	PRT;	536 AA.
ID	Q647H4;		
AC	Q647H4;		
DT	25-OCT-2004 (TrEMBLrel. 28, Created)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Arachin Any-1.		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RP	Yan Y., Wang L., Huang S.;		
RT	"cDNA clone of peanut seed storage protein gene.";		
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY722685; AAU21490.1; -. 30DBF638719AEB78 CRC64;		
SQ	SEQUENCE 536 AA; 61505 MW; 30DBF638719AEB78 CRC64;		
Alignment Scores:			
Pred. No.:	1-02e-163	Length:	536
Score:	2455.00	Matches:	471
Percent Similarity:	93.99%	Conservative:	14
Best Local Similarity:	91.28%	Mismatches:	19
Query Match:	91.23%	Indels:	12
DB:	2	Gaps:	3
US-10-728-323-3 (1-1524) x Q647H4 (1-536)			
QY	1	CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCACGCGCTCAATGCGCAGACCTGAC	60
DB	24	ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnA-gProAep	43
QY	61	AATCGCATTAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACACAGAGGATTC	120

Db	44	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCGCGCGCTCGCCCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCCT	180
Db	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	83
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATATCTTTGGGTTG	240
Db	84	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	103
QY	241	ATATTCCTCGTGTCTTACGACTATGAAGAGCTCACACAAGGTCTGATCTCAG	300
Db	104	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln	123
QY	301	TCCCAAGACACCAAGACCTCCAGGAGAGACCAAGAGCCAAAGCAACGAGATAGT	360
Db	124	SerGlnArgProProArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer	143
QY	361	CACCAGAAGGTGCACCGCTTCGATGAGGGTGATCTCATTCGAGTTCACCGGCTGTGCT	420
Db	144	HisGlnIleValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
QY	421	TTCTGGCTCTACAACGACGACGACGACGACGACGACGACGACGACGACGACGAC	480
Db	164	PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	183
QY	481	AACAAAGACACCAAGCTTGATCATCTCCCGAGAGATTCATTTGGCTGGGAACCGAG	540
Db	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu	203
QY	541	CAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAGCAAGCAAGCAAGCAAGCAAT	600
Db	204	GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr	220
QY	601	AGCCCATACGCCGCGAACTAGCTAGCTAGACAGAGAGCGTGAAATTTAGCCCTCAGGA	660
Db	221	SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly	240
QY	661	CAGCAGCGCGCAGAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	720
Db	241	GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	260
QY	721	AGCGGCTTACGCCGCGAGTCTCTGGAAACAGCCCTCCAGGTTGACACAGACAGATAGT	780
Db	261	SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu	280
QY	781	CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG	840
Db	281	GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly	300
QY	841	GGCCTCAGAACTTTGAGCCCGAGATAGAAAGAGA-----CGTGCAGCAGAA	885
Db	301	GlyLeuArgIleLeuSerProAspArgGlySerArgGlnGlnTyrGluArgProAspGlu	320
QY	886	GAAGAGGAATACGATGAAGATGAATATGAATACGATGAAG-----GATAGA	933
Db	321	GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluGlnHisAspArg	340
QY	934	AGCGGTGGCAGGGAAGCAGAGCGCGGGAATGGTATTGAAGAGAGCATCTGCACGCA	993
Db	341	ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla	360
QY	994	AGTGCTAAAAGAACTTGGTAGAAACAGATCCCTGACATCTACAAACCCCTCAAGCTGGT	1053
Db	361	SerPheIleLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly	380
QY	1054	TCACTCAAACTCCCAACGATCTCAACCTTCTAATACTTAGGTGGCTTGGACCTAGTGCT	1113
Db	381	SerLeuLysThrAlaAsnGluLeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAla	400
QY	1114	GAATATGAATCTTACAGGAATGCAATTTGTCCTCACTACAAACCAACCGCACAC	1173

Db	401	GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis	420
QY	1174	AGCATCATATATCGATTGAGGGACGGCTCAGCTCAAGTCTGTCAGCAGCAACGGCAAC	1233
Db	421	SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp	440
QY	1234	AGAGTTCACGACGAGGAGCTTCAAGAGGTCACGCTTGTGTGTGTGTGTGTGTGTGTGT	1293
Db	441	ArgValPheAspGluGluLeuGlnGlyHisValLeuValValProGlnAsnPheAla	460
QY	1294	GTCGCTGGAAAGTCCAGAGCGGAGAACTTCGATAGCTGCGCATTCAGACAGACTCAAG	1353
Db	461	ValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAspSerArg	480
QY	1354	CCAGCATAGCAACCTCCGCGTGAACCTCGTCATAGATAACCTCGCGAGGAGGTG	1413
Db	481	ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal	500
QY	1414	GTTGCAAAATTCATATGGCTCCAAAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAG	1473
Db	501	ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsnPro	520
QY	1474	TTCAAGTCTTCTCGTCCACGCTCTCAGCAGTCTCCGAGGCGCTGTGGCT	1521
Db	521	PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla	536

RESULT 4

ID	Q647H3	PRELIMINARY;	PRT;	537	AA.
AC	Q647H3;				
DT	25-OCT-2004	(TrEMBLrel. 28, Created)			
DT	25-OCT-2004	(TrEMBLrel. 28, Last sequence update)			
DT	25-OCT-2004	(TrEMBLrel. 28, Last annotation update)			
DE	Arachin Ahv-2.				
OS	Arachis hypogaea (Peanut).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;				
OC	Arachis.				
OX	NCBI_TaxID=3818;				
RN	(1)_TaxID=3818;				
RP	SEQUENCE FROM N.A.				
RA	Yan Y., Wang L., Huang S.;				
RT	"CDNA clone of peanut seed storage protein gene."				
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AV722686; AAU21491.1; -				
SQ	SEQUENCE 537 AA; 61531 MW; 3BDD034DEA159657 CRC64;				

Alignment Scores:

Pred. No.:	1.8e-163	Length:	537
Score:	2451.50	Matches:	474
Percent Similarity:	93.62%	Conservative:	10
Best Local Similarity:	91.68%	Mismatches:	20
Query Match:	91.10%	Indels:	13
DB:	2	Gaps:	4

US-10-728-323-3 (1-1524) x Q647H3 (1-537)

QY	1	CGCGCAGCACCGGAGGAGAACCGTGCAGCTCCAGCGCTCAATGCGCAGAGACCTGAC	60
Db	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	AATCGCATTTGAATTCAGAGCGCGTTTACATTGAGACTTGGAAACCCCAACACGAGGATTC	120
Db	44	AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCGCGCGGTGCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTGTAGGCCT	180
Db	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	83
QY	181	TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATATCTTTGGGTTG	240
Db	84	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	103

```
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCTCACACACAGTGTCTGATCTCAG 300
Db 104 ILePheProGlyCySProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
QY 301 TCCCAAGAGACACCAAGACGTCTCCAAGGAGAGACCAAGAC---CAACAGCAACGAGAT 357
Db 124 SerGlnArgAlaProArgArgPheGlnGlyGluGlnGlnGlnGlnGlnGlnGln 143
QY 358 AGTCACCAAGAGGTGACCGTTTCGATGAGGGTGCATCTCAATGTCAGTTCACCGGTGTT 417
Db 144 SerHisGlnLysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 163
QY 418 GCTTTCGTCTACACGACGACGACACTGATGTTGCTGCTGTTCTTCTTACTGCACAC 477
Db 164 AlaLeuTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 183
QY 478 AACAAACACCAACACGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGACACG 537
Db 184 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHis 203
QY 538 GAGCAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCA 597
Db 204 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 220
QY 598 TATAGCCCATACGCCCGCAAGTACGCTCCTAGACAAAGAGAGCGTGAATTTAGCCCTCGA 657
Db 221 TyrSerProTyrSerProGlnSerGlnProArgGlnGluGluGluGluPheSerProArg 240
QY 658 GGACAGCACAGCGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 717
Db 241 GlyGlnHisSerArgArgGluArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
QY 718 TTCAGCGCTTCACGCCGAGTTCCTGGAAACAAGCTTCAGGTTGACGACAGACAGATA 777
Db 261 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 280
QY 778 GTGCAAAACCTAGAGCGCAGACGAGAGTGAAGAGAGAGGAGCGCATTTGACAGTGAGG 837
Db 281 LeuGlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArg 300
QY 838 GGAGGCTCAGACTCTTGACCCAGATAGAAAGAGA-----CGTGGCGAC 882
Db 301 GlyGlyLeuArgIleLeuSerProAspArgGlyArgGlnGlnTyrGluArgProAsp 320
QY 883 GAAGAAGAGGAATACGATGAAGATGAATATCAATACATCAAGAG-----GAT 930
Db 321 GluGluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluGlnGlnAsp 340
QY 931 AGAAGCGTGGCAGGGGAAGCAGAGCGAGGGGAATGGTATTGAAGAGACGATCTGCACC 990
Db 341 ArgArgArgGlyArgGlySerArgGlyArgGlyAsnGlyIleGluGluThrIleCysThr 360
QY 991 GCAAGTCTAAAGAACATTGTTAGAAACAGATCCCTGCATCTACACCTCAAGCT 1050
Db 361 AlaSerValLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAla 380
QY 1051 GGTTCACCTCAAACTGCCAACGATCTCAACCTCTTAATCTTAGGTGGCTTGGACCTAGT 1110
Db 381 GlySerLeuLysThrAlaAsnAspLeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSer 400
QY 1111 GCTGAATATGGAATCTCTACAGAATGCATTTGTTGCTGCTCACTACAAACACACGCA 1170
Db 401 AlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAla 420
QY 1171 CACAGCATATATATCGATTGAGGGAGCGGCTCAGTGCAGTCACTGTCGACACGACCGC 1230
Db 421 HisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGly 440
QY 1231 AACAGAGTGCACGACGAGCTTCAAGAGGGTCACTGCTTGTGGTGCACAGAACTTC 1290
Db 441 AsnArgValTyrAspGluGluLeuGlnGlnGlyHisValLeuValValProGlnAsnPhe 460
```

```
QY 1291 GCCTGCTCGTGAAGTCCAGAGCGAGAACTTCGAATACGTGGCATTCAAGACAGACTCA 1350
Db 461 AlaValAlaGlyLysSerGlnSerAspAsnPheGluTyrValAlaPheLysThrAspSer 480
QY 1351 AGCCCCAGCATAGCCAACTCGCGGTGAAAACCTCCGTCTATAGATAACCTGCCGAGGAG 1410
Db 481 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 500
QY 1411 GTGGTTGCAAAATTATATGCGCTCCAAAGGAGCAGCAGCAGGCGAGCTTAAGAAACAAC 1470
Db 501 ValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnAsn 520
QY 1471 CCCTTCAAGTCTTCTGCTCCACCGTCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
Db 521 ProPheLysPhePheValProProSerGlnGlnSerLeuGlyAlaValAla 537

RESULT 5
Q8LKN1 PRELIMINARY; PRT; 538 AA.
ID Q8LKN1
AC Q8LKN1
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Allergen Arab3/Arab4
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Viquez O.M., Konan K.N., Dodo H.W.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF510854; AAM46958.1; -.
DR HSSP; P04776; 1UCX.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; RmlC_like_cupin.
DR InterPro; IPR006044; Seedstore_1ls.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 1ISGLOBULIN.
DR PROSITE; PS00305; 1IS_SEED_STORAGE; UNKNOWN_1.
SQ SEQUENCE 538 AA; 61737 MW; 7AABDD59429709E CRC64;

Alignment Scores:
Pred. No.: 6,04e-163 Length: 538
Score: 2444.00 Matches: 471
Percent Similarity: 93.63% Conservative: 14
Best Local Similarity: 90.93% Mismatches: 19
Query Match: 90.82% Indels: 14
DB: 2 Gaps: 4

US-10-728-323-3 (1-1524) x Q8LKN1 (1-538)
QY 1 CGSCAGCAACCGGAGGAGAACCGTCCAGCTTCAGCGCTCAATCGCAGACACCTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGATTGAATCAGAGGCGGTTCATATTGAGACTTGGNACCCCAACACGAGGATTC 120
Db 44 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnProAsnGlnGluPhe 63
QY 121 GAATGCGCGGCTGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 83
QY 181 TTCTACTCCATGCTCCCGAGGAGATCTTCATCCAGCAAGAGGAGGATATCTTTGGGTTG 240
Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCTCGTGTCTCTAGACACTATGAAGAGCTTCACACACAGGTCGTGATCTCAG 300
```

```
Db 104 ILePheProGlyCyPProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
Qy 301 TCCCAAGAGCCACAGAGCTCTCCAGGAGAGACCAAGGCCAACAGCAACGAGATAGT 360
Db 124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnGlnAspSer 143
Qy 361 CACCAGAAAGTGACACCGCTTCGATGAGGGGTGATCTCATTCAGCTGCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
Qy 421 TTCTGGCTCTAACAGCACGACGACTGATGTTGTTGCTGCTTCTTACTGACACCAAC 480
Db 164 PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
Qy 481 AACCAAGCACACCGCTTGATGATCTCCCGAGGATTCATTTGCTGGGACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu 203
Qy 541 CAAGAGTCTTAAAGTACCAGCAACAAGCAGACAAAGCAGACGAAAGACTTACCATAT 600
Db 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
Qy 601 AGCCCATACAGCCGCAAGTCAAGCTAGACAGAGAGCGTGAAATTAGCCCTCGAGGA 660
Db 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
Qy 661 CAGCACAGCCGACAGAACAGCAGCAGACAGAGAGAGAAACCAAGGTGGAACATCTTC 720
Db 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnIlePhe 260
Qy 721 AGCGGCTTCACCGCGAGTCTCTGGAACAAGCTTCAGGTTGACGACAGACAGATAGTG 780
Db 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspArgGlnIleLeu 280
Qy 781 CAAACCTAAGAGCGGAGCCGAGAGTAGTGAAGAGAGAGAGAGGCCATTGTGACAGTGGGA 840
Db 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
Qy 841 GGCTCAGATCTTGACCCAGATAGAAAGAG-----CGTGCAGCAGAA 885
Db 301 GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAspGlu 320
Qy 886 GAAGAGGAATACGATCAAGATGAATATGAATACGATCAAG-----GATAGA 933
Db 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluArgGlnGlnAspArg 340
Qy 934 AGCGTGGCAGGGAGACGAGGAGCGGGGAATGGTATTGAAGAGACGATCTGCACCGCA 993
Db 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
Qy 994 AGTGCTAAAAGAACATTTGTTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGT 1053
Db 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
Qy 1054 TCACTCAAAAGCTGCCACGAT-----CTCAACCTCTTAATACTTAGGTGGCTTGGACCT 1107
Db 381 SerLeuLysThrAlaAsnGluLeuGlnLeuAsnLeuIleLeuArgTrpLeuGlyLeu 400
Qy 1108 AGTGCTGAATATGGAATCTACAGGAATGCATTGTTGCTCCTCACTACACACCAAC 1167
Db 401 SerAlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsn 420
Qy 1168 GCACACAGCATATATATGATTGAGGGAGCGGCTCACGTGCAAGTCTGGGACAGCAAC 1227
Db 421 AlaHisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsn 440
Qy 1228 GGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACGTGCTGTGGTGCACAGAAC 1287
Db 441 GlyAspArgValPheAspGluLeuGlnGlnGlyHisValLeuValValProGlnAsn 460
Qy 1288 TTGCGCTCGCTGGAAAGTCCGAGCGGAGAACTTGGAATACGTGGCATTCAGACAGAC 1347
```

```
Db 461 PheAlaValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAsp 480
Qy 1348 TCAAGSCCCAGCATAGCCAACTCGCGGTGAAAACCTCGTTCATAGATAACCTGCCGGAG 1407
Db 481 SerArgProSerIleIleAlaAsnLeuAlaGlyGluAsnSerPheIleAsnLeuProGlu 500
Qy 1408 GAGGTGGTTGCAAAATTCATATGCGCTCCAAAGGAGCAGCAGGAGGCTTAAAGAACAC 1467
Db 501 GluValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
Qy 1468 AACCCCTTCAAGTCTTCTGTTCCACGCTCTACGAGTCTCCGAGGCGCTTGGCT 1521
Db 521 AsnProPheLysPhePheValProSerGluGlnSerLeuArgAlaValAla 538

RESULT 6
O6T2T4 PRELIMINARY; PRT; 536 AA.
ID O6T2T4
AC O6T2T4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Storage protein.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Yang H.-X., Wang F., Bi Y.-P.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY439332; AAR02860.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR011051; Rmlc_like_cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
SQ SEQUENCE 536 AA; 61498 MW; 138F3C881BC2C7D5 CRC64;

Alignment Scores:
Pred. No.: 2,2e-162 Length: 536
Score: 2436.00 Matches: 468
Percent Similarity: 93.60% Conservative: 15
Best Local Similarity: 90.70% Mismatches: 21
Query Match: 90.52% Indels: 12
DB: 2 Gaps: 3

US-10-728-323-3 (1-1524) x O6T2T4 (1-536)
Qy 1 CGGCAGCAACCGAGAGAACCGGTGCCAGCTTCCAGCGCTCAATCGCGAGAGACTTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
Qy 61 AATCGCATTTGAATCAGAGGGCGGTTCATTTGAGACTTTGGAACCCCAACCAACAGAGATTTC 120
Db 44 AsnArgIleLeuSerGluGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
Qy 121 GAATCGCGCGGCTCGCCCTCTCTCGCTTAGTTCCTCCGCGCCCAAGCCCTTCGTAGGCCT 180
Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 83
Qy 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTTC 240
Db 84 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgAlaTyrPheGlyLeu 103
Qy 241 ATATTCCCTGGTGTCTAGACACATATGAAGAGCCCTCACACACAGGTCTCGATCTCAG 300
Db 104 IlePheLeuGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
```

Qy	301	TCCCAAAGACCAACGACGCTCTCAAGGAGAGAACCAAGCCCAACGACGACGAGATAGT	360
Db	124	SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer	143
Qy	361	CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCACCGGTGGT	420
Db	144	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
Qy	421	TTCTGGCTCTCAACGACACGACACGATGTTGTGTCTGTTCTTCTTACTGACACCAAC	480
Db	164	PheTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	183
Qy	481	AACAAACGACCAACGACTTGATCAGTTCGCCAGGAGATTCATTTGGCTGGGAAACCGAG	540
Db	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
Qy	541	CAAGAGTTCTTAAAGTACCAGCAACAAAGCAGACAGAAAGCAGACCAAGAGCTTACCATAT	600
Db	204	GlnGluPheLeuAlaGlyGlnGln-----GlnSerArgArgArgSerLeuProTyr	222
Qy	601	AGCCCATACACCCGCCAAAGTCACGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	221	SerProTyrSerProGlnThrGlnProIysGlnGluAspArgGluPheSerProArgGly	240
Qy	661	CAGCACAGCCCGCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	720
Db	241	GlnHisGlyArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	260
Qy	721	AGCCGCTTCACGCCGAGTTCCTCGAACACAGCCTTCAGGTTGACGACGACGACGATAGT	780
Db	261	SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu	280
Qy	781	CAAAACCTTAAGAGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	840
Db	281	GlnAsnLeuArgGlyGluAsnGluSerAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	300
Qy	841	GGCTCAGAACTTTGAGCCCAAGTAGAAAGAGA-----CGTCCGACGACAA	885
Db	301	GlyLeuArgIleLeuSerProAspArgGlyIysArgGlnGlnTyrGluArgProAspGlu	320
Qy	886	GAACGGATACGATGAAGATCAATATCAATACCATGACGAG-----GATAGA	933
Db	321	GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluAlaValThrValArgGly	340
Qy	934	AGGCGTGGCAGCGGAAGCAGAGGCGAGGGGGAATGGTATTGAAGAGACGATCTGCACCGCA	993
Db	341	ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla	360
Qy	994	AGTGCTAAAAGAACATTTGTAGAAACAGATCCCTGCATCTACACCCCTCAAGCTGGT	1053
Db	361	SerPheIysIysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly	380
Qy	1054	TCACTCAAAACTGCCACGATCTCAACCTTCTAATCTTAGTGGCTTGGACCTAGTGCT	1113
Db	381	SerLeuIysThrAlaAsnGluLeuAsnLeuLeuIleLeuArgTyrPheGlyLeuSerAla	400
Qy	1114	GAATATGGAATCTCTCAGGAATGTCATTGTTGTGCTCTACTACACCAACCAACGACAC	1173
Db	401	GluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAlaHis	420
Qy	1174	AGCATCATATATCGATTGAGGGGACGGGCTCACGTGCAAGTCGTGGACGACCAACGCAAC	1233
Db	421	SerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp	440
Qy	1234	AGATGTACGACGAGGACTTCAAGAGGGTCACGTGCTTGTGGTGCCACAGAACTTCGCC	1293
Db	441	ArgValPheAspGluGluLeuGlnGlnGlyHisValLeuValValProGlnAsnPheAla	460
Qy	1294	GTCGCTGGAAGTCCCAGAGCGGAACTTCCAATACGTGGCGATTTCACAGACGACTCAAGG	1353
Db	461	ValAlaGlyIysSerGlnSerGluAsnPheGluTyrValAlaPheIysThrAspSerArg	480
Qy	1354	CCAGCATAGCCAACTTCGCCGGTGAAACCTCCGCTCATAGATAACCTGCCGGAGGAGTG	1413

Db 481 ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluGluVal 500
 Qy 1414 GTTGCAAAATCATATGGCGCTCCAAAGGAGGAGCGCAAGGAGCTTAAAGAACAAACACCCC 1473
 Db 501 ValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuIleuysAsnAsnPro 520
 Qy 1474 TTCAAGTTCCTCGTCCACCGCTCTACAGACGTCTCCGAGGGCGTGTGGCT 1521
 Db 521 PheIysPhePheValProProSerGluGlnSerLeuArgAlaValAla 536
 RESULT 7
 Q9SQH7 PRELIMINARY; PRT; 530 AA.
 ID Q9SQH7 AC Q9SQH7
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Glycinin.
 GN Names:Arab4;
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Virginia;
 RX MEDLINE=9406463; PubMed=10474031; DOI=10.1159/000024203;
 RA Kieber-Janke T., Crameri R., Appenzeller U., Schlaak M., Becker W.M.;
 RT "Selective cloning of peanut allergens, including profilin and 2S
 RT albumins, by phage display technology.";
 RL Int. Arch. Allergy Immunol. 119:265-274(1999).
 DR EMBL; AF086821; AAD47382.1; -;
 DR HSP; P04776; I0CX.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR InterPro; IPR011051; RmlC like cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00150; Cupin; 2.
 DR PRINTS; PR00439; IISGLOBULIN.
 SQ SEQUENCE 530 AA; 61011 MW; 864E349A31642A32 CRC64;
 Alignment Scores:
 Pred. No.: 2,48e-161 Length: 530
 Score: 2421.00 Matches: 469
 Percent Similarity: 94.09% Conservative: 9
 Best Local Similarity: 92.32% Mismatches: 29
 Query Match: 89.97% Indels: 2
 DB: 2 Gaps: 0
 US-10-728-323-3 (1-1524) x Q9SQH7 (1-530)
 Qy 1 CGGCAGCAACCGGAGGAGAACGGCGTGCAGTTCAGCGCTCAATGGCAGACCTGAC 60
 Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
 Qy 61 AATCGCATTAATCAGAGGGCGGTTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
 Db 44 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
 Qy 121 GAATGGCGGGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAGCAACGCCCTTCGTAGGCCT 180
 Db 64 GluCysAlaGlyValAlaLeuSerArgLeuValIleuArgArgAsnAlaLeuArgPro 83
 Qy 181 TTCTACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGAGGGGATACTTTGGGTG 240
 Db 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnIleArgGlyTyrPheGlyLeu 103
 Qy 241 ATATTCCCTGGTGTCTCTAGACACTATGAAGACCTTCACACAAGGTCTGCTGATCTCAG 300

104	IlePheProGlyCysProSerThrryrGluIuProAlaGlnGlnGlyArgArgTyrGln	123
301	TCCCAAAGACCAACCAAGACGCTCCAAAGGAGAAGACCAAAAGCCCAACAGCAACGATAGT	360
124	SerGlnArgProProArgArgLeuGlnGluGlnGluAspGlnSerGlnGlnGlnAspSer	143
361	CACCAGAAGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCACCGGTGTGCT	420
144	HisGlnLysValHisArgPheAsnGlnGlyAspLeuIleAlaValProThrGlyValAla	163
421	TTCTGGCTCTACAACGACCGACACATCATGTGTTGCTGCTGTTCTCTTACTGACACCAAC	480
164	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	183
481	AACAACGACACACGACTTGATCAGTCTCCACGAGAGATTCAATTTGGCTGGGAAACGCGAG	540
184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
541	CAAGAGTTCTTAAGGTACCAGCAACAAGCAGACAAAGCAGACGAAGAAGCTTACCATAT	600
204	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	223
601	AGCCCATACAGCCCGCAAGTCAAGCTCAGTACAGCAAGAAGCGTGAATTTAGCCCTCGAGA	660
224	SerProTyrSerProHisSerArgProArgArgGluGluArgGluPheArgProArgGly	243
661	CAGCACAGCCGACAGAACGACGACGACGACCAAGAAGAAGAAACGAGGTGGAACATCTTC	720
244	GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAspGluGlyGlyAsnIlePhe	263
721	AGCGGCTTCACGCGGAGTCTCTGGAACAAGCCTTCCAGGTTGACGACAGACAGATAGT	780
264	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	283
781	CAAAAACCTAAGACGGCGAGACCGACAGCTGACAGAAGAGGGAGGCCATTGACAGTACGGGA	840
284	GlnAsnLeuTrpGlyGluAsnGlnLysSerGluGluGluGlyAlaIleValThrValArgGly	303
841	GGCCTCAGAATCTTTGAGCCGACAGATAGAAAGAGACGTCGCCACGAAAGAGAGAAATACGAT	900
304	GlyLeuArgIleLeuSerProAspGlyThrArgGlyAlaAspGluGluGluTyrAsp	323
901	GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAAAGCAGAGGCAGG	960
324	GluAspGlnTyrGluTyrHisGlnGlnAspGlyArgArgGlyArgGlySerArgGlyGly	343
961	GGGAATGGTATTGAAGAGACGATCTGCACGCCAAGTGTCTAAAAGAACAATTGGTAGAAGC	1020
344	GlyAsnGlyIleGluGluThrIleCysThrAlaCysValLysLysAsnIleGlyGlyAsn	363
1021	AGATCCCTCGACATCTCAACCCCTCAAGC-TGGTTCACTCAAAACTGCCAACGATCTCAA	1079
364	ArgSerProHisIleIleTyrAspProGlnArgTrpPheThrGlnAsnCys-HisAspLeuAs	383
1080	CCTTCTAATACTTAGTGGCTTGGACCTAGTCTGCTCAATATGGAATCTCTACAGGAATGC	1139
383	nLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAl	403
1140	ATTGTTGTGCTCACTACAAACAACACGACACAGCATCATATATTCGATTGAGGGACG	1199
403	aLeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyArg	423
1200	GGCTCAGTCGAAGTCGTGGACACCAACGGCAAACAGAGTGTACGACGAGGAGCTTCAAGA	1259
423	gAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGln	443
1260	GGGTCAAGTCTGTTGTGTGCCACAGAACTTCGCCGTGCTGGAAGTCCCAAGACGAGAA	1319
443	uGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAs	463
1320	CTTCGAATACGTGCATTCAAAGACAGACTCAAGGCCACGACATAGCCAACTCTGCCGCTGA	1379
463	nPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnPheAlaGlyGln	483

QY	1380	AAACTCCGTCATAGATAA	CTGCCGAGGAGGTTGCAAA	TTCATATGGCTCCAAAG	1439
Db	483	uAenSerPheIleAspAsnLeu	ProGluGluValValAlaAsn	SerTyrGlyLeuProAr	503
QY	1440	GGAGCAGGCACAGCGACTT	AAAGAACACCAACCCCTT	CAAGTTCCTCGTCCACCGTCTCA	1499
Db	503	gGluGlnAlaArGglnLeu	LyAsnAsnAsnProPhe	lySPheValProProPheG	523
QY	1500	GCAGTCTCCGAGGCTGTG	GCT 1521		
Db	523	nGlnSerProArGAlaVal	Ala 530		
RESULT 8					
Q61WG5		PRELIMINARY;	PRT;	510 AA.	
ID	Q61WG5	PRELIMINARY;	PRT;	510 AA.	
AC	Q61WG5;				
DT	05-JUL-2004	(T-EMBLrel. 27, Created)			
DT	05-JUL-2004	(T-EMBLrel. 27, Last sequence update)			
DT	05-JUL-2004	(T-EMBLrel. 27, Last annotation update)			
DE	Glycinin (Peanut).				
OS	Arachis hypogaea (Peanut).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;				
OC	Arachis.				
NCBI	NCBI_TaxID=3818;				
RN	[1]	SEQUENCE FROM N.A.			
RP	Kang I.-H., Gallo-Meagher M.;				
RA	Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.				
RL	EMBL; AY618460; AAT39430.1; -.				
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.				
DR	InterPro; IPR006045; Cupin.				
DR	InterPro; IPR011051; RmlC_like_cupin.				
DR	InterPro; IPR006044; Seedstore_11s.				
DR	Pfam; PF00190; Cupin; 2.				
DR	PRINTS; PR00439; 11SGLOBULIN.				
DR	PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.				
FT	NON_TER				
SQ	SEQUENCE 510 AA; 58061 MW; F6569EF13D1A12BD CRC64;				
Alignment Scores:					
Pred. No.:	3-77e-120	Length:	510		
Score:	1834.00	Matches:	364		
Percent Similarity:	78.57%	Conservative:	43		
Best Local Similarity:	70.27%	Mismatches:	71		
Query Match:	68.15%	Indels:	40		
DB:	2	Gaps:	8		
US-10-728-323-3 (1-1524) x Q61WG5 (1-510)					
QY	1	CGCGACGACCGGAGAGACGG	GTGCCAGTTCCAGCGCTCA	TGCGCAGACCTGAC	60
Db	22	ArgGlnGlyGlyGluGluAsn	GluCysGlnPheGlnArg	LeuAsnAlaGlnArgProAsp	41
QY	61	AATCGCATTAATCAGAGCG	CGTTACATTGACACTTGAA	CCCCCAACCAACGAGGTTC	120
Db	42	AsnArgIleGluSerGluGly	GlyTyrIleGluThrTrpAsn	ProAsnAsnGlnGluPhe	61
QY	121	GAATGGCGCGGCTCGCCCT	CTCTCGCTTAGTCTCCGCG	CAACGCCCTTCGTAGCGCT	180
Db	62	GlnCysAlaGlyValAlaLeu	SerArgThrValLeuArg	AsnAlaLeuArgArgPro	81
QY	181	TTCTACTCCAATGCTCCCA	GAGATCTTCATCCAGCAG	GAAGGGGATACTTTGGTTG	240
Db	82	PheTyr-SerAsnAlaProLeu	GluIleTyrValGlnGlnGly	SerGlyTyrPheGlyLeu	101
QY	241	ATATTCCCTGGTTGCTTAG	ACACTATGAAGSCCTCAC	ACACAAGTCGTGCATCTCAG	300
Db	102	IlePheProGlyCysProSer	ThrTyrGluGluProAla	GlnGluGlyArgTyrGln	121
QY	301	TCCCAAGACCCACCAAG	AGCTCTCCAA- - - - -	GGAGAAGACCAAGCCACG	354

```
122 SerGlnYsProSerArgArgPheGlnValGlyGlnAspProSerGlnGlnGln 141
135 GATAGTCACCAAGAGTGCACCGTTTCGATGAGGGTGAATCTCATTCAGTTCACCGGT 414
142 AspSerHisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGly 161
415 GTTCGTTCTTCTGCTCTACAGCAGCAGCAGCAGTGTCTGTTGCTGTTCTCTACTGAC 474
162 ValAlaPheTrpMetTyrAsnAspGluAspValValThrValThrLeuSerAsp 181
475 ACCAAACAACACCAACACAGCTTGATCAGTTCCTCCAGGAGATCAATTTGGCTGGGAAC 534
182 ThrSerSerIleHisAsnGlnLeuAspGlnPheProArgArgPheTyrLeuAlaGlyAsn 201
535 ACGAGCAAGAGTTCCTTAAGTACCAAGCAAAAGCAGACAAAGCAGACCAAGAGCTTA 594
202 GlnGluGlnGluPheLeuArgTyrGlnGlnGlnGlySer----- 215
595 CCATATAGCCATACACCCCGCAAGTCAGCTAGACAGAGAGCGTGAATTTAGCCCT 654
216 -----ArgProHisTyrArgGlnIleSerPro 224
655 CGAGCAGCAGCACCGCGCAGACAGCAGCAGCAAGAAAGAAAGAGTGGAAAC 714
225 ArgValArg-----GlyAspGluGlnGluAsnGluGlySerAsn 237
715 ATCTTCAGCGCTTCACCGCGAGTTCCTGGAACAAGCCTTCAGAGTTGACGACAGACAG 774
238 IlePheSerGlyPheAlaGlnGluPheLeuGlnHisAlaPheGlnVal---AspArgGln 256
775 ATAGTCAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGCGCATTTGACAGTG 834
257 ThrValGluAsnLeuArgGlyGluAsnGluArgGluGlnGlyAlaIleValThrVal 276
835 AGGGAGCGCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTGCC----- 879
277 LysGlyGlyLeuArgIleLeuSerProAspGluAspGluSerSerArgSerProPro 296
880 GACGAAGAAGAGGAATACGATGAAGATGAATATATACGATGAAGAGGATAGAAGCGT 939
297 SerArgArgGluGluPheAspGluAspArg-----SerArgProGlnGlnArg 312
940 GGCAGG-----GGAAGCAGAGCGAGGGGAATGCTATTGAAGAGCAGATCTGCACC 990
313 GlyLysTyrAspGluAsnArgArgGlyTyrLysAsnGlyIleGluGluThrIleCysSer 332
991 GCAAGTCTAAAAGAAACATTGGTAGAAACAGATCCCTCAGATCTACAAACCTCAAGCT 1050
333 AlaSerValLysAsnLeuGlyArgSerSerAsnProAspIleTyrAsnProGlnAla 352
1051 GGTTCACTCAAACTGCCAACGATCTCAACCTTCTAATACTAGGTGGCTTGGACCTAGT 1110
353 GlySerLeuArgSerValAsnGluLeuAspLeuProIleLeuGlyTrpLeuGlyLeuSer 372
1111 GCTGAATATGAATCTCTACAGGAATGCAATTTGTCCTCTCACTACAAACCAAGCA 1170
373 AlaGlnHisGlyThrIleTyrArgAsnAlaMetPheValProHisTyrThrLeuAsnAla 392
1171 CACAGCATATATATGATTGAGGGAGCGGCTCAGCTCAAGTCTGGGACAGCAACGGC 1230
393 HisThrIleValValAlaLeuAsnGlyArgAlaHisValGlnValValAspSerAsnGly 412
1231 AACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTGTTGGTCCACAGAACTTC 1290
413 AsnArgValTyrAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPhe 432
1291 GCGTCTCGTGAAGTCCAGAGCGGAGAACTTCGAATACGTGGATTCAGACAGACTCA 1350
433 AlaValAlaAlaLysAlaGlnSerGluAsnTyrGluTyrLeuAlaPheLysThrAspSer 452
1351 AGCCCAAGCATAGCAACCTCGCGGTGAACCTCCGTATAGTAACTCCGTCAGTACCTCGCGAGGAG 1410
```

```
453 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleIleAspAsnLeuProGluGlu 472
1411 GTGGTTGCAAAATTCATATGCGCTTCCAAAGGAGCAGCAAGGAGCTTAAGAAACAAC 1470
473 ValValAlaAsnSerTyrArgLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 492
1471 CCCTTCAAGTTCTTCGTTCCACCG---TCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
493 ProPheLysPheValProPheAspHisGlnSerMetArgGluValAla 510

RESULT 9
Q647H2 PRELIMINARY; PRT; 484 AA.
ID Q647H2 Q647H2;
AC Q647H2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arachis hvy-3.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP SEQUENCE FROM N.A.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AV722687; AAU21492.1; -.
SQ SEQUENCE 484 AA; 54568 MW; 5A3E950752E89D2D CRC64;

Alignment Scores:
Pred. No.: 2,58e-107 Length: 484
Score: 1651.00 Matches: 337
Percent Similarity: 75.25% Conservative: 46
Best Local Similarity: 66.21% Mismatches: 76
Query Match: 61.35% Indels: 50
DB: 2 Gaps: 8

US-10-728-323-3 (1-1524) x Q647H2 (1-484)
QY 1 CGGCAGCAACCGGAGGAGACCGGTCCAGTTCAGCGCTCAATCGCAGAGACCTGAC 60
DB 24 ArgGlnGlnGlyGluGlnGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTTGAATCAGAGGGCGGTTACATTGAGACTTTGGAACCCCAACCAACAGGAGTTC 120
DB 44 AsnCysIleGluSerGluGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGGTCCCTCTCTCGTTCCTCTAGTCTCTCGTTCAGTTCAGTTCAGTTCAGTTC 180
DB 64 GlnCysAlaGlyValAlaLeuSerArgPheValLeuArgArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGGATCTTTGGGTTC 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleTyrGlnGlySerGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTCGTGTCTCTAGACACTATAGAGCGCTCACACAAAGTCTGTCATCTCAG 300
DB 104 IlePheProGlyCysProGlyThrPheGluGluPro---IleGlnGly---SerGluGln 121
QY 301 TCCCAAGACCAACAGACGTCCTCCAAAGAGAGACCAAGCAAGCAACGACGAGATAGT 360
DB 122 PheGlnArgProSerArgHisPheGlnGlyGlnAspGlnSerGlnArgProLeuAspThr 141
QY 361 CACCAGAGGTGACCGTTTCGATGAGGGTGAATCTCATTCAGTTCCTCCACCGGTGTGCT 420
DB 142 HisGlnLysValHisGlyPheArgGluGlyAspLeuIleAlaValProHisGlyValAla 161
QY 421 TTCTGCTCTACAAACGACCAACGACACTGATGTTGTTGTTGTTCTTCTTCTTCTGACCAAC 480
DB 162 PheTrpIleTyrAsnAspGlnAspThrAspValValAlaIleSerValLeuHisThrAsn 181
```



```
QY 481 AACAAACGACACAGCTTCAATCCAGGAGATTCAATTTGGCTGGGAACACGAG 540
   ::: |||||
Db 182 SerLeuHisGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyLeuGlnGlu 201
   |||||
QY 541 CAAAGATTCTTAAGTTACCAAGCAACAAAGC---AGACAAAGCAGACGAAGAAGCTTACCA 597
   |||||
Db 202 GlnGluPheLeuArgTyrGlnGlnArgSerGlyArgGlnSer----- 215
   |||||
QY 598 TATAGCCCATACAGCCCGCAAGTCAGCTAGACNAGAGAGCGTGAATTTAGCCCTCGA 657
   |||::: |||||
Db 216 -----ProLeuGlyGluGlnGlu----- 222
   |||||
QY 658 GGACAGCACAGCCGACAGAACGAGCAGACGAAGAACAAACGAAGGTGGAACATC 717
   |||||
Db 223 -----GlnGluGlnGluGlnGluGlyAsnVal 232
   |||||
QY 718 TTCAGCGCTTCAGCCGAGGATCTCTGGAAACAAGCTTCAGGTTCAGCAGACAGACATA 777
   |||||
Db 233 PheSerGlyPheSerThrGluPheLeuSerHisGlyPheGlnValAsnGlu---AspIle 251
   |||||
QY 778 GTGCNAAACCTAAGAGCGGAGACCGAGAGTCAAGACAGAGGAGGCCATTGTGACAGTGA 837
   |||||
Db 252 ValArgAsnLeuArgGlyGluAsnGluArgGluGlnGlyAlaIleValThrValLys 271
   |||||
QY 838 GGAGGCTCAGAAATCTTGAGCCAGATAGAAAGAGAGAGCTGCCAGCAAGAAAGAGGAATAC 897
   |||||
Db 272 GlyGlyLeuSerIleLeuValPro----- 280
   |||||
QY 898 GATGAAGATGAATATGATGATGATGAAGATAGAAAGCGTGGCAGGGAACAGAGGC 957
   |||||
Db 281 -----GluTyrArgGlnSerTyrGlnGlnProGlyArgGlyAspLys 295
   |||||
QY 958 AGGGGGAATGTTATGAGAGAGCATCTGCACCGCAGTCTAAAGACATCTGTAGA 1017
   |||||
Db 296 PheAsnAsnGlyIleGluGluThrIleCysThrAlaThrValLysMetAsnIleGlyLys 315
   |||||
QY 1018 AACAGATCCCTGCATCTACACCTCAAGCTGTTCTCACTCAAACTGCAACAGTCTC 1077
   |||||
Db 316 SerThrSerAlaAspIleTyrAsnProGlnAlaGlySerValArgThrValAsnGluLeu 335
   |||||
QY 1078 AACCTTCTAATACTAGTGGCTGGACCTAGTGTGTAATATGAAATCTCTACAGAAAT 1137
   |||||
Db 336 AspLeuProIleLeuAsnArgLeuGlyLeuSerAlaGluTyrGlySerIleHisArgAsp 355
   |||||
QY 1138 GCATTGTTCTCGTCACTACACCAACCAAGCAGCAGCATCATATATCATTCAGGGA 1197
   |||||
Db 356 AlaMetPheValProHisTyrAsnMetAsnAlaAsnSerMetIleTyrAlaLeuHisGly 375
   |||||
QY 1198 CGGCTCACTGCAAGTCGTGGACAGCAACGCAACAGAGTGTACACAGGAGCTTCAA 1257
   |||||
Db 376 GlyAlaHisValGlnValValAspCysAsnGlyAsnArgValPheAspGluGluLeuGln 395
   |||||
QY 1258 GAGGTCACTGCTGTGTGGTGCACAGAACTTCGCGTGCCTGGAAAGTCCAGAGCGAG 1317
   |||||
Db 396 GluGlyGlnSerLeuValValProGlnAsnPheAlaValAlaLysSerGlnSerGlu 415
   |||||
QY 1318 AACTTCGAATACGTGGCATTCACAGACACTCAAGGCCAGCATACCCACCTCGCGGT 1377
   |||||
Db 416 HisPheLeuTyrValAlaPheLysThrAsnSerArgAlaSerIleSerAsnLeuAlaGly 435
   |||||
QY 1378 GAAAACTCCGCTCATAGATAACCTGCCGAGAGGTGTTGCCAAATTCATATGCGCTCCAA 1437
   |||||
Db 436 LysAsnSerTyrMetTyrAsnLeuProGluAspValValAlaAsnSerTyrGlyLeuGln 455
   |||||
QY 1438 AGGAGCAGCAGGACGCTTAAGAACAAACAAACCTTCAGTTCCTGTTCCACCG--- 1494
   |||||
Db 456 TyrGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheThrThrPheLeuValProGln 475
   |||||
QY 1495 TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
   |||||
Db 476 AspSerGlnMetIleArgThrValAla 484
```

```
RESULT 10
GLC3_SOYBN STANDARD; PRT; 481 AA.
AC P11828;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycinin G3 precursor (Contains: Glycinin A subunit; Glycinin B
DE subunit).
GN Name-GY3;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]_TaxID=3847;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296500; PubMed=2740231;
RA Cho T.-J., Nielsen N.C.;
RT "The glycinin G3 gene from soybean.";
RL Nucleic Acids Res. 17:4388-4388(1989).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=92393391; PubMed=2485233;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X15123; CAA33217.1; -
DR PIR; S04605; S04605.
DR HSP; P04776; IFXZ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmC like cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
KW Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 296 Glycinin A subunit.
FT CHAIN 297 476 Glycinin B subunit.
FT PROPEP 477 481
FT DISULFID 107 303 Interchain (By similarity).
SQ SEQUENCE 481 AA; 54241 MW; 5F3C3148DF6241A7 CRC64;
Alignment Scores:
Pred. No.: 7,33e-97 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 1 Gaps: 7
US-10-728-323-3 (1-1524) x GLC3_SOYBN (1-481)
QY 1 CGGCAGCAACCGAGGAGAACCGGTCCAGTTCACGGCCTCAATCGCGAGACCTGAC 60
```



```
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
Qy 61 AATCGCATTTGAATCAGAGGGCGGTTCATTGAGACTTTGGAAACCCCAACACAGGAGTTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
Qy 121 GAATGCGCGCGCGTCTCTCGCTTCCTAGTCTCCGCGCAACGCGCTTCGTAGCGCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
Qy 181 TTCTACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGCAAGCGGATATTGGGTTG 240
Db 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
Qy 241 ATATTCCCTGTTTCTTAGACACTATGAAGAGCTTCACACAAAGTCGTGCGATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
Qy 301 TCCCAAGACCAACCAAGACGCTCTCAAGGAGAGACCAAGCAACCAAGCAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
Qy 361 CACCAGAGGTGACCGTTTCGATGAGGTGATCTCATTTGAGTTCACCGCGTGTGCT 420
Db 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
Qy 421 TTCTGGCTCTACACGACCAACGACGACTGATGTTGTTCTGTTCTTCTTACTGACCAAC 480
Db 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
Qy 481 AACCAACACACACGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACAGGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
Qy 541 CAAGAGTTCTTAAGGTACCAAGCAACAAAGACAGACAAAGCAGACGACGAAGCTTACCATAT 600
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
Qy 601 AGCCCATACACCGCCGAAAGTCAGCTTAGACAAGAGCGGTGAATTTAGCCCTCGAGGA 660
Db 200 -----Gly 200
Qy 661 CAGCAGACCGCAGACAGAACGACGAGGACAGAAAGCAAGAAACGAGGTGAACATCTTC 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 220
Qy 721 AGCGGCTTCACGCGGAGTTCCTCGAACAAGCCTTCAGGTTGACGACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
Qy 781 CAAAACCTTAAGACGCGAGACGAGAGTGAAGAAGAGGAGGCCATTGTGACAGTGAAGGGA 840
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluLysGlyAlaIleValThrValLysGly 259
Qy 841 GGCCTCAGATCTTGACCCCA-----GATAGAAGAGACGTGCCGACGAGAAGAG 891
Db 260 GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu 279
Qy 892 GAATPACCATGAAGATGAATATGATACGATGAGGATAGAAGCGTGGCAGGGAAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
Qy 952 AGAGGCGAGGGGAATGGTATTGAGAGACGATCTGCAACCCCAAGTGTCAAAAAGAACATT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
Qy 1012 GGTAGAACAGATCCCTCGACATCTCAACCTCAAGCTGGTTCACCTCAAAACATGCCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
Qy 1072 GATCTCAACCTCTTAATACATTAGTGGCTTGGACCTAGTCTGAAATATGGAATCTCTAC 1131
|||||
```

```
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
Qy 1132 AGGAATGCATTTGTTGCTCTACTACAACCAACGACACACAGCATATATATCGATTG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
Qy 1192 AGGGGACGGGCTCACCTGCAAGTCGTGGACAGCAACGCGACAGAGTGTACGACGAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
Qy 1252 CTTCAAGAGGGTCACTGCTTGTGGTCCACAGAACTTCCCGCTCGCTGGAAGTCCCGAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
Qy 1312 AGCGAAGACTTCAATACGTGGCATTCACAGACAGACTCAAGGCCCGACATAGCCAACTC 1371
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
Qy 1372 GCCGGTGAACACTCCCTCATAGATACCTGCCGAGGAGGTGCTTGCAAATTCATATGCG 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
Qy 1432 CTCCAAGGAGGACGAGCAAGCGAGCTTAAGAAACAACAACCCCTTCAAGTTCTTCGTTCGA 1491
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
Qy 1492 CCCTCTCAGACGCTCTCCGAGGGCTGGCT 1521
Db 472 ProLysGluSerGlnArgValValAla 481
RESULT 11
GLC2_SOYBN STANDARD; PRT; 485 AA.
ID 20-405; P04121; P04348; P04349;
DC 20-MAR-1987 (Rel. 04, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Glycinin G2 precursor [Contains: Glycinin A2 subunit; Glycinin B1A subunit].
GN Names=G2;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=92393391; PubMed=2485233;
RA Nielsen N.C., Dickinson C.D., Cho T.J., Thanh V.H., Scallan B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296499; PubMed=2740230;
RA Thanh V.H., Turner N.E., Nielsen N.C.;
RT "The glycinin Gy2 gene from soybean.";
RL Nucleic Acids Res. 17:4387-4387(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Shirotseurunoko;
RA Utsumi S., Kim C.S., Kohno M., Kito M.;
RT "Polymorphism and expression of cDNAs encoding glycinin subunits.";
RL Agric. Biol. Chem. 51:3267-3273(1987).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88040439; PubMed=3671077;
RA Fukazawa C., Momma T., Higuchi W., Uda K.;
RT "Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit precursor of soybean.";
RL Nucleic Acids Res. 15:8117-8117(1987).
```


Db 238 ArgAenLeuGlnGlyGluAenGluGluAspSerGlyAlaIleValThrValValysGly 257
Qy 841 GGCTCTAGATCTTGGACCCAGATAGAAAGAGAGCTGCCGCGAAGAGAGGAATACGAT 900
Db 258 GlyLeuArgValThrAlaProAlaMetArgIysProGlnGlnGluGluAspAsp 277
Qy 901 GAAGATGAA-----TATGAATACCATCAAGAGGATAGAACGCGTGGCAGGGGA 948
Db 278 GluGluGlnProGlnCysValGluThrAspIysGlyCysGlnArgGlnSerLysArg 297
Qy 949 AGCAGAGCGAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGGCTAAAGAAC 1008
Db 298 SerArg-----AenGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAen 314
Qy 1009 ATTGGTGAACAGATCCCTGACATCTACACCTCAAGCTGGTTCACCTCAAACTGCC 1068
Db 315 IleGlyGlnAenSerProAspIleTyraenProGlnAlaGlySerIleThrAla 334
Qy 1069 AACGATCTCAACCTTCTAATCTTAGTGGCTTGGACCTAGTCTGTAATGGAATCTC 1128
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuIysLeuSerAlaGlnTyrglySerLeu 354
Qy 1129 TACAGGAATCATTTGTTGCTGCTACTACCAACACGACAGCAGCATCATATATCGA 1188
Db 355 ArgIysAenAlaMetPheValProHisTyThrLeuAenAlaAenSerIleIleTyAla 374
Qy 1189 TTGAGGGAGCGGCTCAGCTGCAAGTGTGAGACAGCAACGGCAACAGAGTGTACGACGAG 1248
Db 375 LeuAenGlyArgAlaLeuValGlnValAlaAenCysAenGlyGluArgValPheAspGly 394
Qy 1249 GAGCTTCAAGAGGCTCAGCTGCTGTGTGTCGCCAGAACTTCGCGCTGGGAACTGCC 1308
Db 395 GluLeuGlnGlyGlyValLeuIleValProGlnAenPheAlaValAlaAlaLysSer 414
Qy 1309 CAGAGCGAGAACTTCGAATACGTGGCTTCAAGACAGACTCAAGCCGACGATAGCCAAAC 1368
Db 415 GlnSerAenAenPheGluTyValSerPheIysThrAenAspArgProSerIleGlyAen 434
Qy 1369 CTCGCCGGTGAACCTCCGTATAGATAAATCTGCCGAGAGGTTGGCAATTCATAT 1428
Db 435 LeuAlaGlyAlaAenSerLeuLeuAenAlaLeuProGluGluValIleGlnHisThrPhe 454
Qy 1429 GGCTCTCAAGGGAGCAGGAGGACGCTTAAAGAACACACCCCTTCAAGTCTTCGTT 1488
Db 455 AenLeuIysSerGlnGlnAlaArgGlnValIysAenAenAenProPheSerPheLeuVal 474
Qy 1489 CCACCGTCTCAGCTCTCCGAGGCGTGTGCT 1521
Db 475 ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 12

GLC1 SOYBN STANDARD; PRT; 495 AA.
AC P04776;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Glycinin G1 precursor [Contains: Glycinin A1a subunit; Glycinin BX subunit].
GN Name=G1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OC NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=92393391; PubMed=2485233;
RA Nielsen N.C., Dickinson C.D., Cho T.J., Thanh V.H., Scallion B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328(1989).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296498; PubMed=2740229;
RA Sims T.L., Goldberg R.B.;
RT "The glycinin Gyl gene from soybean.";
RL Nucleic Acids Res. 17:4386-4386(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Bonminori;
RX MEDLINE=86041867; PubMed=2997720;
RA Negoro T., Momma T., Fukazawa C.;
RT "A cDNA clone encoding a glycinin A1a subunit precursor of soybean.";
RL Nucleic Acids Res. 13:6719-6731(1985).
RN [4]
RP SEQUENCE FROM N.A.
RC Utsunomi S., Kohno M., Mori T., Kito M.;
RT "An alternate cDNA encoding glycinin A1a Bx subunit.";
RL J. Agric. Food Chem. 35:210-214(1987).
RN [5]
RP SEQUENCE FROM N.A.
RC Urade R., Nakatani H., Nakano C.;
RT "mRNA of soybean proglycinin A1aB1b subunit.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -I- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.
CC -I- PTM: The precursor is post-translational processed to form a covalently linked A1a-BX subunit complex.
CC -I- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36686; AAA33966.1; -;
DR EMBL; X15121; CAA33215.1; -;
DR EMBL; X02985; CAA26723.1; -;
DR EMBL; AB113349; BAC78522.1; -;
DR PIR; A23497; FWSY62.
DR PIR; S10851; S10851.
DR PDB; 1EXZ; X-ray; A/B/C=20-495.
DR PDB; 1UCX; X-ray; A/B/C=20-495.
DR PDB; 1UDI; X-ray; A/B/C=20-495.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR011051; RmlC like cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; 1.
KW 3D-structure; Multigene family; Seed storage protein; Signal.
FT SIGNAL 1 19 Glycinin A1a subunit.
FT CHAIN 20 306 Glycinin BX subunit.
FT PROPEP 307 310 Interchain (By similarity).
FT CHAIN 311 490 D -> G (in Ref. 3).
FT PROPEP 491 495 P -> S (in Ref. 3).
FT DISULFID 107 317 F -> S (in Ref. 3).
FT CONFLICT 42 42 E -> G (in Ref. 3).
FT CONFLICT 108 108
FT CONFLICT 136 136
FT CONFLICT 360 360
FT TURN 30 31
FT STRAND 39 39
FT STRAND 43 47
FT TURN 48 49
FT TURN 50 54

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	811	30.1	492	4	US-09-462-720-2	Sequence 2, Appli
2	293	10.9	141	4	US-09-645-593-12	Sequence 12, Appl
3	227.5	8.5	85	4	US-09-645-593-10	Sequence 10, Appl
4	215	8.0	165	4	US-09-645-593-11	Sequence 11, Appl
5	184.5	6.9	489	4	US-09-424-283-3	Sequence 3, Appl
6	173.5	6.4	96	4	US-09-645-593-9	Sequence 9, Appl
7	172.5	6.4	626	4	US-09-106-872A-4	Sequence 4, Appl
8	167	6.2	524	4	US-09-424-283-1	Sequence 1, Appl
9	160	5.9	448	4	US-09-323-195A-18	Sequence 18, Appl
10	158	5.9	523	4	US-09-323-195A-17	Sequence 17, Appl
11	157.5	5.9	335	4	US-09-106-872A-17	Sequence 17, Appl
12	156.5	5.8	587	1	US-07-955-905A-23	Sequence 23, Appl

QY	52	AGACCTGACAATCGCATTTGAATCATCAGAGCGCGGTACATTGAGACTTGGAAACCCCAACAC	111
DB	45	GluproSerPheArgPheProSerGluAlaGlyLeuThrGluPheTrpAspSerAsnAsn	64
QY	112	CAGAGATTTCGAATCGCGCGCGTGCCTCTCTCGTTAGTCCCTCCGCGCAACGCCCTT	171
DB	65	ProGluPheGlyCysAlaGlyValGluPheGluArgAenThrValGlnProGlyGlyLeu	84
QY	172	CGTAGCGTTTCTACTCCAATGCCTCCACGAGATCTTCATCCAGCAAGGAGGAGTAC	231
DB	85	ArgLeuProHisIstYrSerAsnValProLysPheValTyrValValGluGlyThrGlyVal	104
QY	232	TTTGGTGTGATTTCCCTGGTTTCTAGACACTATGAAGAGCCTCACACAAGGTCGT	291
DB	105	GlnGlyThrValIleProGlyCysAlaGluThrPheGlu-----SerGlnGlyGlu	121
QY	292	CGA-----TCTCAGTCCCAAGACCACCAAGAGCTCTCCAAGGAGAGACCA-----	339
DB	122	SerPheTrpGlyGlnGlnGlnProGlyLysGlyGlnGluGlyGlnGlnGlySer	141
QY	340	-----AGCCAAACAGCAACGA-----GATAGTACCAGAAAGTGCACCGT	378
DB	142	LysGlyGlyGlnGluGlyArgArgGlnArgPheProAspArgHisGlnLysLeuArgArg	161
QY	379	TTCGATGAGGTGATCTCATTTGCAGTTCCTCCACCGGTGTGCTTCTTGCTCTACAACGAC	438
DB	162	PheGlnLysGlyAspValLeuIleLeuLeuProGlyPheThrGlnTrpThrTyrAsnAsp	181
QY	439	CACGACACTGATGTGTGCTTCTTCTTACTGACCAACAACAAGCAACACAGCTT	498
DB	182	GlyAspValProLeuValThrValAlaLeuLeuAspValAlaAsnGluAlaAsnGlnLeu	201
QY	499	GATCAGTTTCCCAGGAGATTCAAATTTGGCTGGGAACACGGAGCAAGAGTTCTTTAAGGTAC	558
DB	202	AspLeuGlnSerLysPhePheLeuAlaGlyAsnProGlnGlnGlyGlyGlyLysGlu	221
QY	559	CAGCAACAAGACAGACAAGACAGACAGAAAGACTTTACCATATAGCCCATACAGCCGCA	618
DB	222	GlyHisGlnGlyGlnGlnGlnHisArg-----	231
QY	619	AGTCAGCCTACACAAGACGCGTGAATTTAGCCTCGAGACAGCACAGCCGCAGAGAA	678
DB	231	-----	231
QY	679	CGAGCAGACAAGAAGAAAAACGAGGTGAAACATCTTTACGCGCTTCACGCCGAG	738
DB	232	-----AsnIlePheSerGlyPheAspAspGln	240
QY	739	TTCTGTGGAACAGCCTTCCAGTTGCACACAGATAGTGTCAAAACCTTAAGAGCGCAG	798
DB	241	LeuLeuAlaAspAlaPheAsnVal--AspLeuLysIleIleGlnLysLeuLysGly---	258
QY	799	ACCGAGTGAAGAGAGCGGAGCCATTGTGTGACGTGAGGGA-----GGCTCAGAACTTGT	855
DB	259	--ProLysAspGlnArgGlySer-----ThrValArgAlaGluLysLeuGlnLeuPhe	275
QY	856	AGCCAGATAGAAAGACAGCGTCCGACGACGAAGAGGAATACGATGAATATATGAA	915
DB	276	LeuProGluTyrSerGluGlnValGlnGlnProGlnGln-----	288
QY	916	TACGATGAAGAGGATAGAACGCGTGGCAGGGGAAACGAGCAGCGGGGAATGGTATTGAA	975
DB	289	---GlnGlnGlnGlnGlnHisGlyValGlyArgGlyTrpArgSerAsnGlyLeuGlu	307
QY	976	GAGACGATCTGCACGCGAAGTGTATAAAGAACATTTGGTAGAAACAGATCCCTTGACATC	1035
DB	308	GluThrLeuCysThrValLysLeuSerGluAsnIleGlyLeuProGlnGluAlaAspVal	327
QY	1036	TACAACCTCAAGCTGTTCTCAAAACTGCCAACGATCTCAACCTTCTAATATTAGG	1095
DB	328	PheAsnProArgAlaGlyArgIleThrValAsnSerGlnLysIleProIleLeuSer	347
QY	1096	TGGCTTGGACCTAGTGTGTAATATGGAAATCTTACAGGAATGATTTGTTGCTGCAC	1155

Db	348	SerLeuGlnLeuSerAlaGluArgGlyPheLeuTy-SerAsnAlaIlePheAlaProHis	367
Qy	1156	TACAACACCAACGACACACATATATCGAATTCAGGAGCGGGCTCACGTGCAAGTC	1215
Db	368	TrpAsnIleAsnAlaHisAsnAlaLeuTy-ValIleArgGlyAsnAlaArgIleGlnVal	387
Qy	1216	GTGCACAGCAACCGCACAGAGTGTACGACGAGGAGCTTCAAGAGGCTCACGTGTTGTG	1275
Db	388	ValAspHisIysGlyAsnLysValPheAspAspGluValLysGlnGlyLeuLeuIleIle	407
Qy	1276	GTGCCACAGAACTTCGCGCTCGTGGAAAGTCCAGAGCGGAGAACTTCGNAATACGTGGCA	1335
Db	408	ValProGlnTyrPheAlaValIleLysLysAlaGlyAsnGlnGlyPheGluTyrValAla	427
Qy	1336	TTCAAGACACACTCAAGCGCCACATAGCCAACTCGCCGGTGAACACTCCGTCATAGAT	1395
Db	428	PheLysThrAsnAspAsnAlaMetIleAsnProLeuValGlyArgLeuSerAlaPheArg	447
Qy	1396	AACTCTCGGAGAGGTGGTGTCAAAATTCATATGGCTCCAAAGGAGCAGGCAAGCAG	1455
Db	448	AlaIleProGluGluValLeuArgSerSerPheGlnIleSerGluGluAlaGluGlu	467
Qy	1456	CTTAAGAACCAACCCCTCAAGTTCTTCGTTCCACCGTCTCAGCAGTCT	1506
Db	468	LeuLysTyrGlyArgGlnGluArgLeuLeuLeu-----SerGluGlnSer	482
RESULT 2			
US-09-645-593-12			
; Sequence 12, Application US/09645593			
; Patent No. 6777591			
; GENERAL INFORMATION:			
; APPLICANT: Chaudhary, Sarita			
; APPLICANT: van Rooijen, Gijs			
; APPLICANT: Moloney, Maurice			
; APPLICANT: Singh, Surinder			
; TITLE OF INVENTION: Flax Seed Specific Promoters			
; FILE REFERENCE: 9369-151			
; CURRENT APPLICATION NUMBER: US/09/645,593			
; CURRENT FILING DATE: 2000-08-25			
; PRIOR APPLICATION NUMBER: US 60/151,044			
; PRIOR FILING DATE: 1999-08-27			
; PRIOR APPLICATION NUMBER: US 60/161,722			
; PRIOR FILING DATE: 1999-10-27			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 12			
; LENGTH: 141			
; TYPE: PRT			
; ORGANISM: Linum usitatissimum			
US-09-645-593-12			
Alignment Scores:			
Pred. No.:	1,21e-19	Length:	141
Score:	293.00	Matches:	55
Percent Similarity:	72.73%	Conservative:	25
Best Local Similarity:	50.00%	Mismatches:	30
Query Match:	10.89%	Indels:	0
DB:	4	Gaps:	0
US-10-728-323-3 (1-1524) x US-09-645-593-12 (1-141)			
Qy	1138	GCATTGTTTTCCTCCTACATCAACACCAACGACACAGCATCATATATCGATTGAGGGA	1197
Db	2	AlaIleArgLeuProHisTrpAsnIleAsnAlaHisSerIleValTyrAlaIleArgGly	21
Qy	1198	CGGCTCTCACCTGCAAGTCGTGGAGCAACCGGCAACAGAGTGTACGACGAGGAGCTTCAA	1257
Db	22	GlnAlaArgValGlnIleValAsnGluGluGlyAsnSerValPheAspGlyValLeuGln	41
Qy	1258	GAGGGTCACTGTTGTGGTGCACAAACTTCGCGCTCGCTCGGAAGTCCCAAGCGCAG	1317
Db	42	GluGlyGlnValValThrValProGlnAsnPheAlaValValLysArgSerGlnSerGlu	61

Db 99 -----GlyGlyArgGlyGlyArgTyrTyrSerAsnGlyValGluGluThrPhe 115
Qy 985 TGCACCCCAAGTGTCTAAAAAGAACATTGGTAGAACAGATCCCTCGACATCTACAAACCT 1044
Db 116 CysSerMetArgLeuIleGluAsnIleGlyAspProSerArgAlaAspIlePheThrPro 135
Qy 1045 CAAGCTGGTTCACATAAACTGCCAACGATCTCAACCTTCTAATATCTTAGTGGCTTGA 1104
Db 136 GluAlaGlyArgValArgSerLeuAsnSerHisAsnLeuProValLeuGluThrPheGln 155
Qy 1105 CCTAGTCTGAATATGGAATCTCTAC 1131
Db 156 LeuSerAlaGluArgGlyValLeuTyr 164
RESULT 5
US-09-424-283-3
; Sequence 3, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-3
Alignment Scores:
Pred. No.: 6.51e-09 Length: 489
Score: 184.50 Matches: 99
Percent Similarity: 34.19% Conservative: 87
Best Local Similarity: 18.20% Mismatches: 187
Query Match: 6.86% Indels: 171
DB: 4 Gaps: 17
US-10-728-323-3 (1-1524) x US-09-424-283-3 (1-489)
Qy 1 CGCAGCAACCGGAGGAGACGGTCCAGTTCAGCGGCTCAATGGCAGACCT--- 57
Db 80 ArgGluLysGluGluGluHisGlnGluGlnHisGluGluGluAspGluAsnProTyr 99
Qy 58 -----GACAATCGCATTGAATCAGAGCGGCGTTACATTGAG 93
Db 100 ValPheGluGluAspLysAspPheSerThrArgValGluThrGluGlySerIleArg 119
Qy 94 ACTTGGAAACCCCAACACAGGAGTTCGAA---TGGCGCGGCTCGCCCTCTCTCGCTTA 150
Db 120 ValLeuLysLysPheThrGluLysSerLysLeuLeuGlnGlyIleGluAsnPheArgLeu 139
Qy 151 GTCTCT-----CGCCCAACGCCCTTCGTAGGCTTCTACTCCAATGCTCCCCAG 201
Db 140 AlaIleLeuGluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluVal 159
Qy 202 GAGATCTTCACTCCAGCAAGGAGGGGATCTTTGGGTTCATATTCCTGCTGCTCCTAGA 261
Db 160 ValLeuPheAsnIleLysGlyArgAlaValLeuGlyLeuVal----- 173
Qy 262 CACTATGAAGAGCTCACACAAAGGTGTCGATCTCAGTCCCAAGACCAACCAAGACGT 321
Db 173 ----- 173
Qy 322 CTCCAGGAGAGACCAACCAACAGACGAGATAGTCACCAAGGTGCACCGCTTC 381
Db 174 -----ArgGluSerGluThrGluLysIleThrLeu 183

Qy 382 GATGAGGGTGATCTCATTTGCAGTTCACACCGGTGTTGCTTTCTGCGCTACACACCCAC 441
Db 184 GluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArgAsp 203
Qy 442 GACACTGATGTTGTTGCT 501
Db 204 GluAsnGluLysLeuLeuLeuAlaMetLeu----- 213
Qy 502 CAGTTCCTCCCGAGGATTCATTTGGCTGGGAACACGGAGCAAGAGTTCCTTAAGGTACCAG 561
Db 213 ----- 213
Qy 562 CAACAAAGCAGCAAAAGCAGACGAAGAGCTTACCATATAGCCCATACAGCCGCCAAAGT 621
Db 214 -----HisIleProValSer 218
Qy 622 CAGCCTAGACAAGAGAGCGTGAAATTTACCTCGAGGACAGCACAGCGCGCAGAGAACA 681
Db 219 ThrProGlyLysPheGluGluPheGlyProGlyIleArg----- 232
Qy 682 GCAGGACAGAAGAAACCAAGGTGGAACATCTTCAGCGGCTTCACGCGGAGTTC 741
Db 233 -----AspProGluSerValLeuSerAlaPheSerTrpAsnVal 245
Qy 742 CTGGAACAAGCCTTCAGGTTGACGACACACAGATAGTCAAAACCTTAAGAGCGCAGACC 801
Db 246 LeuGlnAlaLeuGlnThrPro-----LysGlyLysLeu 257
Qy 802 GAG-----AGTGAAGAAGAGGAGGCGCATTTGTACAGTG---AGGGAGGCGCTC 846
Db 258 GluArgLeuPheAsnGlnGlnAsnGluGlySerIlePheLysIleSerArgGluArgVal 277
Qy 847 AGAATCTTCAGCCCGACATAGAAGAGACGTGCCGACGAGAGAGAGAAATACGATGAAGAT 906
Db 278 ArgAlaLeuAlaProThrLysLys-----SerSer 287
Qy 907 GAATATGAATACGATGAAGAGGATAGAAGCGTGGCAGGGGAAGCAGAGCGAGGGGAAT 966
Db 288 TrpTrpProPheGlyGlySerLys----- 296
Qy 967 GGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACATTTGGTAGAACAGATCC 1026
Db 297 -----AlaGlnPheAsnIlePheSerLysArg--- 305
Qy 1027 CCTGACATCTACAAACCTCAAGCTGTTTCACTCAAAACCTGCCAACGATCTCAACCTCTA 1086
Db 306 ProThrPheSerAsnGlyTyrGlyArgLeuThrGluValGlyProAspAspGluLysSer 325
Qy 1087 ATACTTAGTGGCTTGGACCTAGTGTGAATATGAAATCTCTACAGGAATGCATTGTTT 1146
Db 326 TrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSerMetSer 345
Qy 1147 GTCGCTCACTACACACCAACCGCACACATCATATATCGATTGAGGGGAGCGGCTCAC 1206
Db 346 ThrIleHisThrAsnSerHisAlaThrIleAlaLeuValMetAspGlyArgGlyHis 365
Qy 1207 GTGCAAGTC-----GTGGACAGCAACCGCAACAGAGGTGTACACAG----- 1248
Db 366 LeuGlnIleSerCysProHisMetSerSerArgSerAspSerLysHisAspLysSerSer 385
Qy 1249 -----GAGCTTCAAGAGGCTCACGTGCTTGTGGTCCACAG 1284
Db 386 ProSerTyrHisArgIleSerAlaAspLeuLysProGlyMetValPheValProPro 405
Qy 1285 AACTTCGCC---GTGCTGGAAGTCCCGACCGGAGAACTTCGAATACGTGCATTCAG 1341
Db 406 GlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuIleCysPheGlu 425
Qy 1342 ACAGACTCAAGCCCGACCATAGCC---AACCTCGCGGTGAAACTCCCTCATATA----- 1392
Db 426 ValAsnValArgAspAsnLysLysPheThrPheAlaGlyLysAspAsnIleValSerSer 445

Oy 1393 --GATACCTGCGGAGGAGGTGGTTGCAATTCATATGCGCTCCAAAGGAGCAGGCA 1449
 Db 446 LeuaspnValAlaLysGluLeuAlaPheAsnTyrProSerGluMetValAsnGlyVal 465
 Oy 1450 AGGCAGCTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAGCAGTCTCCG 1509
 Db 466 SerGluArgLysGluSerLeuPhePheProPheGluLeuProSerGluGluArgGlyArg 485
 Oy 1510 AGGCTGTGGCT 1521
 Db 486 ArgAlaValAla 489

RESULT 6

US-09-645-593-9
 ; Sequence 9, Application US/09645593
 ; Patent No. 677591
 ; GENERAL INFORMATION:
 ; APPLICANT: Chaudhary, Sarita
 ; APPLICANT: van Rooijen, Gijb
 ; APPLICANT: Moloney, Maurice
 ; APPLICANT: Singh, Surinder
 ; TITLE OF INVENTION: Flax Seed Specific Promoters
 ; FILE REFERENCE: 9369-151
 ; CURRENT APPLICATION NUMBER: US/09/645,593
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR FILING DATE: 1999-08-27
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; TYPE: PRT
 ; LENGTH: 96
 ; ORGANISM: Linum usitatissimum
 ; US-09-645-593-9

Alignment Scores:
 Pred. No.: 3.24e-08 Length: 96
 Score: 173.50 Matches: 35
 Percent Similarity: 63.51% Conservative: 12
 Best Local Similarity: 47.30% Mismatches: 24
 Query Match: 6.45% Indels: 3
 DB: 4 Gaps: 2

US-10-728-323-3 (1-1524) x US-09-645-593-9 (1-96)

Oy 1 CGGCAGCAACGGAGGAG--AACGCTGCGCAGTTCCAGCGCTCAATGCGCAGACCT 57
 Db 25 ArgGlnGlnPheGlnGlnGlyAsnGluCysGlnIleAspArgIleAspAlaSerGluPro 44
 Oy 58 GACAAATCGCAATGATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACACCGAGGAG 117
 Db 45 AspLysThrIleGlnAlaGluAlaGly-----GluValTrpAspGlnAsnArgGlnGln 62
 Oy 118 TTCGAATCGCGCGCTGCGCTCTCTCGTTAGTCTCCGCGCGCAACGCCCTTCGTAGG 177
 Db 63 PheGlnCysAlaGlyAlaValAlaValArgThrIleGluProLysGlyLeuLeu 82
 Oy 178 CTTTCTACTCCCAATGCTCCCGAGGAGATCTTCATCCAGAA 219
 Db 83 ProPheYrSerAsnThrProGlnLeuIleTyrIleValGln 96

RESULT 7

US-09-106-872A-4
 ; Sequence 4, Application US/09106872A
 ; Patent No. 6486311
 ; GENERAL INFORMATION:
 ; APPLICANT: Burks Jr., A. Wesley
 ; APPLICANT: Stanley, J. Steven
 ; APPLICANT: Cockrell, Gael
 ; APPLICANT: King, Nina E.
 ; APPLICANT: Sampson, Hugh A.

; APPLICANT: Helm, Ricki M.
 ; APPLICANT: Bannion, Gary A.
 ; TITLE OF INVENTION: Peanut Allergens and Methods
 ; FILE REFERENCE: HS 103 CIP
 ; CURRENT APPLICATION NUMBER: US/09/106,872A
 ; PRIOR FILING DATE: 1999-06-29
 ; PRIOR FILING DATE: 1996-09-23
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: Arachis hypogaea
 ; FEATURE:
 ; OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 1
 ; OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 2
 ; OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 3
 ; OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 4
 ; OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 5
 ; OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 6
 ; OTHER INFORMATION: Amino Acids 123-132 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 7
 ; OTHER INFORMATION: Amino Acids 134-143 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 8
 ; OTHER INFORMATION: Amino Acids 143-152 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 9
 ; OTHER INFORMATION: Amino Acids 294-303 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 10
 ; OTHER INFORMATION: Amino Acids 311-320 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 11
 ; OTHER INFORMATION: Amino Acids 325-334 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 12
 ; OTHER INFORMATION: Amino Acids 344-353 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 13
 ; OTHER INFORMATION: Amino Acids 393-402 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 14
 ; OTHER INFORMATION: Amino Acids 409-418 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 15
 ; OTHER INFORMATION: Amino Acids 461-470 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 16
 ; OTHER INFORMATION: Amino Acids 498-507 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 17
 ; OTHER INFORMATION: Amino Acids 525-534 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 18
 ; OTHER INFORMATION: Amino Acids 539-548 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 19
 ; OTHER INFORMATION: Amino Acids 551-560 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 20
 ; OTHER INFORMATION: Amino Acids 559-568 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 21
 ; OTHER INFORMATION: Amino Acids 578-587 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 22
 ; OTHER INFORMATION: Amino Acids 597-606 are Ara H 1 binding epitope,
 ; OTHER INFORMATION: peptide 23
 ; OTHER INFORMATION: peptide 23
 ; US-09-106-872A-4

Alignment Scores:

Pred. No.: 1.06e-07 Length: 626
 Score: 172.50 Matches: 117
 Percent Similarity: 35.22% Conservative: 82
 Best Local Similarity: 20.71% Mismatches: 204
 Query Match: 6.41% Indels: 163
 DB: 4 Gaps: 25

US-10-728-323-3 (1-1524) x US-09-106-872A-4 (1-626)

```

QY 4 CAGCAACCGGAGGAGAACGGGTGCTCCAGTTCAGCGCCTCAATCGCGAGACCTGACAAT 63
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
130 ArgGlnProArgLysAsp-----TipArgArgProSerHisGlnGlnPro---Arg 145
QY 64 CGATTGGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGAGTTCGAA 123
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
146 LysIleArgProGluGly-----ArgGluGlyGluGlnGluTrpGly 159
QY 124 TGGCGCGCGTCCCTCTCTCGCTTAGCTCTCGCGCGCAACGCCCTTCGTAGCCCTTC 183
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
160 ThrProGlySerHisValArgGluGluThrSerArgAsnAsn-ProPhe----- 175
QY 184 TACTCCAATCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTGATA 243
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
176 -----TyrPheProSerArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgI1 193
QY 244 TTCCTGGTGTCTAGACACTATGAAGAC-----CTCACACACAA 285
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
193 eArgValLeuGlnArgPheAspGlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisAr 213
QY 286 GGTGCTCGATCTCAGTCCCAAA-----GACCACCAAGACGTCTC 324
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
213 GIIeValGlnIleGluAlaLysProAsnThrLeuValLeuProLysHisAlaAspAlaAs 233
QY 325 -----CAAGGAGAGAACCAAGCCACACCAACGAGATAGTCACCA 365
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
233 pAsnIleLeuValIleGlnGlnAlaThrValThrValAlaAsnGlyAsnAsnAr 253
QY 366 GAAGTGCACCGTTTCATGAGGGTGNATCTCATTTGAGTTCCACCGGTGCTCTTCG 425
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
253 gLysSerPheAsnLeuAspGluGlyHisAlaLeuArgIleProSerGlyPheIleSerTy 273
QY 426 GCTCTACAACGACGACGACTGATGTGTGCTGTTCTCT---CTTACTGACACCAACAA 482
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
273 rIleLeuAsnArgHisAspAsnGlnAsnLeuArgValAlaIleSerMetProValAs 293
QY 483 CAACGACAAACGACTGATCAGTTGCCAGGAGATTCAATTTGGCTGGGAAACGAGCA 542
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
293 nThrProGlyGlnPheGluAspPhe-----PheProAlaSerSerArgAspGlnSe 310
QY 543 AGAGTTCTTAAGTTACGACCAACAAAGCAGACCAAGCAGACGAGAAAGCTTACCATA 602
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
310 rSerTyrLeu-----GlnGluPheSerArgAsnThrLeu----- 321
QY 603 CCATACAGCCGCAAAAGTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACA 662
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
322 -----GluAlaAlaPheAsnAlaGluPheAs 330
QY 663 GCACAGCCGCAGA-----GAACGAGCAGGACAGACGAGAGTGAAGAA- 710
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
330 nGluIleArgArgValLeuLeuGluGluAsnAlaGlyGlyGluGlnGluGluArgGly-- 349
QY 711 AAACATCTTCAGCGGCTTCAGCGCGGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAG 770
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
349 ----- 349
QY 771 ACAGATAGTCAAAAACCTAAGAGCGAGACCGAGAGTGAAGAA-----GAGGGAGCCAT 824
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
350 -----GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValI1 365
QY 825 TGTGACAGTGAAGGAGGCGCTCAGAATCTTGAAGCCAGATAGAAAGAGACGTCGCCACGA 884
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
365 eValIleVal----- 368
QY 885 AGAAGAGGATACGATGAGATGAATATGATAGATGAAGAGGATAGAGGGTGGCAG 944
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
369 -SerLysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLysLysGlySe 388
QY 945 GCGAAGCAGAGGAGGCGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCAAAA 1004
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
388 rGluGluGluGly-----AspIleThrAsnProIleAsnLeuArgGluGluPr 405
QY 1005 GAACATTGGTGAAGAAC-----AGATCCCTGACATCTACAAACCTCA 1046

```

```

Db 405 oAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProG1 425
QY 1047 ACCTGGTTCTACCTCAAAACTGCCAACGATCTCAACCTTCTTAATACTTAGGTGGCTGGACC 1106
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
425 n-----LeuGlnAspLeuAspMetMetLeu----- 433
QY 1107 TAGTGTCTGAATATGAAATCTCTACAGGAATGCATTGTTTGTGCTCTACTACAACACCAA 1166
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
434 -----ThrCysValGluIleLysGluGlyAlaLeuMetLeuProHisPheAsnSerLy 451
QY 1167 CGCACACACATCATATATCGATTGAGGGGACGGGCTCACGTCGAAGTCGTG----- 1218
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
451 sAlaMetValIleValValAsnLysGlyThrGlyAsnLeuGluLeuValAlaValAr 471
QY 1219 -----GACAGCAA 1226
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
471 gLysGluGlnGlnArgGlyArgGluGluGluAspGluAspGluGluGluG1 491
QY 1227 CGGCAAC-----AGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTGTTGT 1274
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
491 uGlySerAsnArgGluValArgArgTyrThrAlaArgLeuLysGluGlyAspValPheI1 511
QY 1275 GTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAGCGAGAACTTCGAATACGTGCG 1334
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
511 eMetProAlaAlaHisProValAlaIleAsnAlaSerSerGlu---LeuHisLeuLeuG1 530
QY 1335 ATTC-----AAGACAGACTCAAGGCCAGCATGCCAACCTCGCGGTGAAACCTCCGT 1388
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
530 yPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsnVa 550
QY 1389 CATAGATAACCTCGCGAGGAGTGTGTCGAAATTCATATGCGCTCCAAAGGGAGCAGGC 1448
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
550 lIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGlnVa 570
QY 1449 AAGCGAGCTTAAGAACAAACACCCCTTCAAGTTCCTCGTT-----CCACCGTCTCA 1499
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
570 lGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAlaArgProGlnSerG1 590
QY 1500 G---CAGTCTCCG 1509
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
590 nSerGlnSerPro 594

RESULT 8
US-09-424-283-1
; Sequence 1, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.
; TITLE OF INVENTION: Sucrose binding proteins
; FILE REFERENCE: 4630-50206
; CURRENT APPLICATION NUMBER: US/09/424,283
; CURRENT FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: PCT/US98/10465
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: US 60/047,568
; PRIOR FILING DATE: 1997-05-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
US-09-424-283-1

Alignment Scores:
Pred. No.: 3,27e-07 Length: 524
Score: 167.00 Matches: 104
Percent Similarity: 33.98% Conservative: 89
Best Local Similarity: 18.31% Mismatches: 179
Query Match: 6.21% Indels: 196
DB: 4 Gaps: 22

```



```
US-10-728-323-3 (1-1524) x US-09-424-283-1 (1-524)
QY 1 CGGACGACCGGAGGAGAACCGCTGCAGCTTCAGCGCCCTCAATGCGCAGACCT--- 57
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
92 ArgGluArgGluGluGluGlnGlnGlnHisGluGlnAep---GluAsnProTyr 110
QY 58 -----GACANTCGATTCAATCAGAGGCGGTTACATTGAG 93
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
111 IlePheGluGluAspLysAspPheGluThrArgValGluThrGluGlyGlyArg 130
QY 94 ACTTGGAAACCCCAACACAGGAGGTTCCAA---TGCGCCGCGCTCGCCTCTCTCGCTTA 150
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
131 ValLeuLysLysPheThrGluLysSerLysLeuGlnGlyLleGluAsnPheArgLeu 150
QY 151 GTCCTC-----CGCGCAACGCCCTTCGTAGGCTTTCTACTCCANTGCTCCCGAG 201
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
151 AlaIleLeuGluAlaArgAlaHisThrPheValSerProArgHisPheAspSerGluVal 170
QY 202 GAGATCTTCATCCAGCAGGAGGAGGATCTTTGGGTTGATATTTCCCTGGTGTCTCTAGA 261
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
171 ValPhePheAsnIleLysGlyArgAlaValLeuGlyLeuVal----- 184
QY 262 CACTATGAAGAGCTCACACAAAGGTCTCGATCTCAGTCCCAAGACCCACCAAGACGT 321
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
185 -----SerGluSerGluThr----- 189
QY 322 CTCACAGGAGAGACCAAGCAACAGACAGATAGTCAACAGAGGTGCACCGTTTC 381
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
190 -----GluLysIle-----ThrLeu 194
QY 382 GATGAGGTTGATCTCATTGCGATTCCACCGGTGTCTTCTTGGCTCTACAGACCAC 441
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
195 GluProGlyAspMetIleHisIleProAlaGlyThrProLeuTyrIleValAsnArgAsp 214
QY 442 GACACTGATGTGTGCTGTTCTCTACTGACACCAACCAACGACCAACCGCTTGAT 501
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
215 GluAsnAspLysLeuPheLeuAlaMetLeu----- 224
QY 502 CAGTTCGCCAGGAGATTCAATTTGGCTGGGAACACAGGCAAGAGTTCTTAAGGTACCAG 561
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
224 ----- 224
QY 562 CAACAAAGCAGACAAAGCAGCAGAGAGCTTACCATATAGCCCATACAGCCCGCAAGT 621
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
225 -----HisIleProValSer-----ValSer 231
QY 622 CAGCTAGACAAGAGCGTGAATTTAGCCCTCGAGCAGACGACACGCGCAGAGAACA 681
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
232 ThrProGlyLysPheGluGluPhePheAlaProGlyArg----- 245
QY 682 GCAGGACAAGAAGAAAACGAAGGTGGAACATCTTCAGCGGCTTCACGCCGAGTTC 741
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
246 -----AspProGluSerValLeuSerAlaPheSerTrpAsnVal 258
QY 742 CTGGAACAAGCCTTCAGGTTGACGACAGACAGATAGTGCAAACTTAAGAGGGGAGACC 801
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
259 LeuGlnAlaAlaLeuGlnThrProLysGlyLys-----LeuGluAsnValPhe 274
QY 802 GAGAGTCAAGAAGAGGAGGACCATGTGCACAGTG---AGGGAGGCGCTCAGAACTTCGAGC 858
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
275 AspGlnGlnAsnGluGlySerIlePheArgIleSerArgGlnValArgAlaLeuAla 294
QY 859 CCAGATAGAAAGAGACGTCGCCGACGAAGAGGAATACGATGAAGATGAATATGATATAC 918
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
295 ProThrLysLys-----SerSerTrpTrpProPhe 304
QY 919 GATGAAGAGATAGA-----AGGCGTGGCAGGGAAGCAGAGGC 957
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
305 GlyGlyGluSerLysProGlnPheAsnIlePheSerLysArgProThrIleSerAsnGly 324
QY 958 AGGGGGAATGGTATTGATGAGAGACCATCTGACCCCGCAAGTGTCTAAAGACATTTGGTAGA 1017
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
325 TyrGlyArgLeuThrGlu----- 330
```

```
1018 AACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACTCAAACTGCCAAGCATCTC 1077
331 -----ValGlyProAspAspAspGluLysSerTrpLeuGlnArgLeu 344
1078 AACCTTCTTAATCTAGTGGCTGACCTAGTGTCTGAATATGAAATCTCTACAGGAT 1137
345 AsnLeuMetLeu-----ThrPheThrAsnIleThrGlnArg 356
1138 GCATTGTTTTCGCTCACTACACACCAACGACACAGCATCATATATCGATTGAGGGA 1197
357 SerMetSerThrIleHisTyrAsnSerHisAlaThrLysIleAlaLeuValIleAspGly 376
1198 CGGCTCACGTGCAAGTC-----GTGACAGCAACGCGCAACAGGTGTAC 1242
377 ArgGlyHisLeuGlnIleSerCysProHisMetSerSerArgSerHisSerLysHis 396
1243 GACGAG-----GAGCTTCAAGAGGGTCACTGCTT 1272
397 AspLysSerSerProSerTyrHisArgIleSerSerAspLysLysProGlyMetValPhe 416
1273 GTGCTGCCACAGAACTTCCGCC---GTGCTGGAAAGTCCAGAGCGAGCAAGCAGCTT 1329
417 ValValProProGlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuMet 436
1330 GTGGCATTTCAAGACAGACTCAAGGCCCGCAGCATAGCC---AACCTCGCGGTGAAACTCC 1386
437 IleCysPheGluValAsnAlaArgAspAsnLysLysPheThrPheAlaGlyLysAspAsn 456
1387 GTCATAGAT-----AACCTGCGGAGGAGGTG 1413
457 IleValSerSerLeuAspAsnValAlaLysGluLeuAlaPheAsnTyrProSerGluMet 476
1414 GTTGCAAAATTCATATATGCGCTCCAAAGGAGGAGCAGGCAAGGAGCAGCTT----- 1458
477 ValAsnGlyValPheLeuLeuGlnArgPheLeuGluArgLysLeuIleGlyArgLeuTyr 496
1459 -----AAGAACAACAACCCCTTCAAGTTCTTCGTTCCACCGTCT 1497
497 HisLeuProHisLysAspArgLysGluSerPhePheProPheGluLeuProArgGlu 516
1498 CAGCAGTCTCCGAGGCTGTGGCT 1521
517 GluArgGlyArgAlaAspAla 524

RESULT 9
US-09-323-195A-18
; Sequence 18, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Perrera, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Picea glauca
US-09-323-195A-18

Alignment Scores: 1.43e-06 Length: 448
Pred. No.: 160.00 Matches: 103
Score: 33.59% Conservative: 70
Percent Similarity: 20.00% Mismatches: 172
Best Local Similarity: 5.95% Indels: 170
Query Match: 4 Gaps: 20
DB:
```

```
US-10-728-323-3 (1-1524) x US-09-323-195A-18 (1-448)
QY 4 CAGCAACGGAGGAGACGGTGCAGTTCAGCGCCTCAATGGCAGAGACCTGACAAT 63
Db 46 GLUGluArgGLUGluAsnProTyrValPhe-----HisSerAspSerPheArgThr 62
QY 64 CGCATTGAATCAGAGGCGGTTCATTAGACTTGGAAACCCCAACACACAGGAGTTGCGAA 123
Db 63 ArgAlaSerSerGLuAlaGLyGLuIleArgAlaLeu---ProAsnPheGLyGLuValSer 81
QY 124 -----TGCCTGGGTGGCGCTCTCTCGTTAGTC-----CTCCGCCCAAGGCC 168
Db 82 GLuLeuLeuGLUGlyIleArgLysPheArgValThrCysIleGLuMetLysProAsnThr 101
QY 169 CTTTCGTAGGCTTTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGCAAGGGGA 228
Db 102 ValMetLeuProHisTyrIleAspAlaThrTrpIleLeuTyrValThrArgGLyArgGLy 121
QY 229 TACTTTGGGTGTGATTTCCCTGGTTGTCTAGACACTATGAAGAGCCTCACACACAAGGT 288
Db 122 TyrIleAlaTyrVal-----126
QY 289 CGTCGATCTCAGTCCCAAGACCAACAGACGCTCTCAAGGAGAGACCAACCAACAG 348
Db 127 -----HisGLnAsnGLuLeuValLysArgLysLeu-----136
QY 349 CAACGAGATAGTCACCAAGAGGTGCACCGTTTCGATGAGGTCATCTCATTGCAGTCC 408
Db 137 -----GLUGlyAspValPheGLyValPro 145
QY 409 ACCGGTGTCTTTCTGGCTCTCAACGACGACACTGATGTTGTGTCTCTCTT 468
Db 146 SerGLyHisThrPheTyrLeuValAsnAspHisAsnThrLeuArgIleAla---164
QY 469 ACTGACACCAACCAACGACCAACGAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCT 528
Db 164 -----164
QY 529 GGGAAACCGGACGAGAGTTCTTAAGTACCAGCAACAAAGCAGACAAAGCAGCAAGA 588
Db 164 -----164
QY 589 AGCTTACATATAGCCCATACAGCCGCAAGTCAGCTAGACAAAGAGCGTGTAATTT 648
Db 165 -----SerLeuValArgProValSerThrValArgGLyGLuTyrGlnProPhe 180
QY 649 AGCCTCGAGGACGACCGCCGAGAGACGAGCAGGACCAAGAAAGAAAGAGGT 708
Db 181 TyrValAlaGLy-----GlyArgAsnProGln-----189
QY 709 GGAACATCTTCAGCGCTTCACGCGGAGTTCCTGGAACAAGCCTTCAGGTTGACGAC 768
Db 190 -----ThrValTyrSerAlaPheSerAspValLeuGLuAlaAlaPheAsnThrAsnVal 208
QY 769 AGACAGATAGTGCAAAACCTAAGAGGCGAGACCGAGAGTGAAGAAGAGGAGGAGCATTTGTG 828
Db 209 GlnGLnLeuGLuArgIlePheGLyGLy-----HisLysSerGLyValIleIle 224
QY 829 ACAGTGAGGGGAGCCTCAGAACTTTGAGCCAGATAGAAAGAGAGCGTCCGACGAAGAA 888
Db 225 -----HisAlaAsnGLuGlu 229
QY 889 GAGGAATACGATGATGAATATGAATACGATGAAGAGATAGAGGCGTGCAGGGGA 948
Db 230 -----GlnIleArgGLuMetMetArgLysArgGLy-----Phe 240
QY 949 AGCAGAGCGGAGGGGAATGATTGAAGAGACGATCTGCACCCGCAAGTGTCTAAAGAAAC 1008
Db 241 SerAlaGLySerMetSerAlaProGLuHis-----ProLysProPhe 254
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACAACCCCTCAGCTGGTTCACTCAAAACTGCC 1068
Db 1009 -----1068
```

```
255 AsnLeuArgAsnGlnLysProAspPheGluAsn---GluAsnGLyArgPheThrIleAla 273
1069 AACGATCTCAACCTTCTTAATACTAGTGGCTTGGACCTAGTCTGAATATGGAATCTC 1128
Db 274 GlyProLysAsnTyrProPheLeuAspAlaLeuAspValSerValGLyLeuAlaAspLeu 293
QY 1129 TACAGGAATGCATTTGTCGCTCACTAACACACACACAGCAGCATCATATATCGA 1188
Db 294 AsnProGLySerMetThrAlaProSerLeuAsnSerLysSerThrSerIleGLyIleVal 313
QY 1189 TTGAGGGGACGGCTCACGTGCAAGTCGTG-----1218
Db 314 ThrAsnGLyGLuGLyArgIleGluMetAlaCysProHisLeuGLyGlnHisGLyTrpSer 333
QY 1219 -----GACACCAACGGCAACAGAGTGTACACGAGGAGCTT 1254
Db 334 SerProArgGLuArgGLyAspGlnAspIleThrTyrGlnArgValTrp---AlaLysLeu 352
QY 1255 CAAGAGGTCACGTGCTTGTGTGTCACAGAACTTGCCTGCTGGA---AAGTCCCAAG 1311
Db 353 ArgThrGLySerValTyrIleValProAlaGLyHisProIleThrGluIleAlaSerThr 372
QY 1312 AGCGAGAACTTCGAATACGTGCATTCAGACAGACTCAAGGCCAGCATAGCCAC---1368
Db 373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGLyAsnGLuArgGlnPhe 392
QY 1369 CTCGCGGTGAAAACCTCCGTATAGTACCTCGCGAGAGGTGGTTGCAAAATTCATAT 1428
Db 393 LeuAlaGLyLysAsnAsnValLeuAsnThrLeuGLuArgGLuIle-----407
QY 1429 GGCCTCCAAAGGAGCAGCGCAAGCGAGCTTAAGAAACAACAACCC 1473
Db 408 -----ArgGlnLeuSerPheAsnValPro 415

RESULT 10
US-09-323-195A-17
; Sequence 17, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Ferreira, Ranjan
; TITLE OF INVENTION: VITILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 17
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Pinus taeda
US-09-323-195A-17

Alignment Scores:
Pred. No.: 2,41e-06 Length: 523
Score: 158.00 Matches: 107
Percent Similarity: 32.67% Conservative: 74
Best Local Similarity: 19.31% Mismatches: 190
Query Match: 5.87% Indels: 183
DB: 4 Gaps: 20

US-10-728-323-3 (1-1524) x US-09-323-195A-17 (1-523)
QY 10 CCGAGGAGGAACGGTGCAGTTCAGCGCCTCAATGCCAGAGAGCCTGACAAAT-----63
Db 37 ProGLuAspHisGLyArgGLyHisGlnArgArgGLuGLuArgGLuGLuAsnProTyr 56
QY 64 -----CGCATTTGAATCAGAGCGGCTTACATTGAGACT 96
Db 57 ValPheHisSerAspArgPheArgMetArgAlaSerSerAspAlaGLyGLuIleArgAla 76
```

```
QY 97 TGAACCCCAACCAACAGGAGTTCGAA-----TGCGCGCGCTGCGCCTCTCTCGCTTA 150
Db 77 Leu---ProAsnGlyGluAlaSerGluLeuLeuGluGlyLeuSerLysTyrArgVal 95
QY 151 GTC-----CTCCGCGCAACAGCCCTTCGTAGGCGCTTCTACTCCAATGCTCCCCAG 201
Db 96 ThrCysIleGluMetArgProAsnThrValMetLeuProHisTyrLeuAspAlaThrTrp 115
QY 202 GAGATCTTCATCCAGCAAGAGGAGTACTTTGGGTGTATATTCCTCGTGTGTCCTAGA 261
Db 116 IleLeuTyrValThrGlyGlyArgGlyTyrIleAlaTyrVal----- 129
QY 262 CACTATGAAGAGCTCACACAAAGTTCGTGATCTCAGTCCCAAGACCAACCAAGACGT 321
Db 130 -----HisGlnAsnGluLeuValLysArgLys 138
QY 322 CTCAAGAGAGAACCAAGCCAAAGCCACAGACAGATAGTCACCAAGAGGTGCACCGTTTC 381
Db 139 Leu----- 139
QY 382 GATGAGGTGATCTCATTGTCAGTTCCACCGGTGTTGCTTTCTGGCTCTCAACGACAC 441
Db 140 GluGluGlyAspValPheGlyValProSerGlyHisThrPheTyrLeuValAsnAsp 159
QY 442 GACACTGATGTTGTTGCTGTTCTCTTACTGACACCAACCAACACGACCAACCGCTTGAT 501
Db 160 AspHisAsnSerLeuArgIleThr----- 167
QY 502 CAGTTCGCCAGGAGATTCAATTTGGCTGGGAACACAGGCAAGAGTTCTTAAGTACCAG 561
Db 168 -----SerLeuLeuArgThrVal 173
QY 562 CAACAAAGCAGACAAAGACGACGAAGAAGTTACATATAGCCCATACAGCCGCCAAAGT 621
Db 174 SerThrMetArgGlyGlu-----TyrGluProTyr----- 183
QY 622 CAGCCTAGACGAAGAGCGGTGAATTTAGCCCTCGAGGACGACACCGCAGAGAACA 681
Db 184 -----TyrValAla 186
QY 682 GCAGGCAAGAAGAAAGAAAGAGGTGGAAACATCTTCAGCGGCTTCAGCGCGAGTTC 741
Db 187 GlyGlyArgAsnProGlu-----ThrValTyrSerAlaPheSerAspAspVal 202
QY 742 CTGGAACACAGCTTCCAGGTGTGACGACAGACAGATAGTGCAAACCTTAAGAGGCGAGAC 801
Db 203 LeuGluAlaAlaPheAsnThrAsn----- 210
QY 802 GAGAGTGAAGAGGAGGCGCATTTGTGCAGTGAAGGAGGCGCTCAGAACTTCGAGCCCA 861
Db 211 -----ValIleGluAlaArgThrHisPheProVal----- 220
QY 862 GATAGAAAGAGACGTGCCGAGAGAGAGAAATACGATGAAGATGAATATGAATACGAT 921
Db 221 -----HisIleGluArgGluSerTyrSerMetAla 230
QY 922 GAAGAGATAGAAGCGGTGGCAGGGAAGCAGAGCGAGGGGAATGTATTTGAAGAGACG 981
Db 231 AsnGluGluGlnIleArg---GluMetLeuArgLysArgGlyPheSerAlaGluSerMet 249
QY 982 ATCTGCACCGCAAGTGTCTAAAGAACATTTGGTAGAACAGATCCCTCGCATCTACAAC 1041
Db 250 SerAlaSerGluHisProLysProPheAsnLeuArgAsnGlnLysProAspPheGluAsn 269
QY 1042 CCTCAAGCTGGTTCACCTCAAACTGCCAACGATCTCAACCTCTTAATCTAGTGGCTT 1101
Db 270 AspAsnGlyArgPheThrArgAlaGlyProAsnGluAsnProLeu---LeuAspAlaVal 288
QY 1102 GGACCTAGTCTGAATATGAAATCTCTACAGGAATGCATTTGTTGCTGCTCACTACAAC 1161
Db 289 AspValThrAlaGlyPheGlyValLeuAsnProGlyThrMetThrAlaProSerHisAsn 308
QY 1162 ACCAACCCACACATCATATATCGATTGAGGGGACGGCTCACGTGCAAGTCGTG--- 1218
```

```
Db 309 ThrLysAlaThrSerIleAlaIleValThrGlnGlyGluGlyArgIleGluMetAlaCys 328
QY 1219 -----GACAGCAAC 1227
Db 329 ProHisLeuGlyGlnHisGlyTrpSerSerArgArgGluLysGlyAspGlnGluIleAsn 348
QY 1228 GGCACACAGAGTGTACGACGAGGAGCTTCAAGAGGTCACGCTGCTGTGGTGGCCACAGAAC 1287
Db 349 TyrGlnArgVal---ArgAlaArgLeuArgThrGlyThrValTyrValValProAlaGly 367
QY 1288 TTCGCGCTGCTCGAAAGTCCAGAGCGAG---AACTTCGAATACGTGGCATTCAAGACA 1344
Db 368 HisProIleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTrpPheAspIle 387
QY 1345 GACTCAAGGCCACGATAGCCAAAC---CTCCGCGGTGAAACCTCCGTCATAGATAACCTG 1401
Db 388 AsnThrSerGlyAsnGluArgGlnPheLeuAlaGlyLysTyrAsnValLeuGlnThrLeu 407
QY 1402 CCGAGAGAGTGTGTCAAATTCATATGGCTCTCAAAAG---GACCAG 1446
Db 408 GluLysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluVal 427
QY 1447 GCAAGGCGAGCTTAAGAAC-----AACAAAC 1470
Db 428 LeuArgArgGlnLysAspGlnValIleLeuArgGlyProGlnMetGlnArgArgAspGlu 447
QY 1471 CCCTTCAAGTTCCTTCGTTCCACGCTCTCAGCAGTCTCTCCGA 1510
Db 448 -ProArgSerSerSerIleHisAlaIleIleAlaAlaArg 460

RESULT 11
US-09-106-872A-17
; Sequence 17, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-17

Alignment Scores:
Pred. No.: 2,14e-06 Length: 335
Score: 157.50 Matches: 55
Percent Similarity: 42.04% Conservative: 48
Best Local Similarity: 22.45% Mismatches: 81
Query Match: 5.85% Indels: 61
DB: 4 Gaps: 9

US-10-728-323-3 (1-1524) x US-09-106-872A-17 (1-335)
QY 31 TTCACGCGCTCAATCGCAGAGACCTGACAATCGCATTCGATTAATCAGAGCGGTACATT 90
Db 122 PheGlyLysLeuPheGluValLysProAspLys----- 133
QY 91 GAGACTTGGNACCCCAACCAACAGGAGTTCGAATGCGCGGCTCGCCCTCTCTCGCTTA 150
Db 91 GAGACTTGGNACCCCAACCAACAGGAGTTCGAATGCGCGGCTCGCCCTCTCTCGCTTA 150
```



```
QY 664 CACAGCCGACAGAACGAGCAGGACGAAGAAAGAAACGAAGGTGGAAACATCTTTCAGC 723
Db   :|||
QY 265 ProGlyLysTyrGluLeuPheProAlaGlyAsnAsnLysProGluSerTyrTyrGly 284
Db   :|||
QY 724 GGTTCACGCCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGAC-----AGACAG 774
Db   :|||
QY 285 AlaPheSerTyrGluValLeuGluThrValPheAsnThrGlnArgGluLysLeuGluGlu 304
Db   :|||
QY 775 ATAGTCAAAACCTAGAGCCGAGACCGGAGAGTGAAGAGAGGAGGAGGCCATTTGACAGTG 834
Db   :|||
QY 305 IleLeuGluGluGlnArgGlyGlnLysArgGlnGlnGlyGln----- 318
QY 835 AGCGGAGCGCTCAGATCTTTGAGCCCGCAGATAGAAAGAGAGCTGCCGACGAAGAAGAGAA 894
Db   :|||
QY 319 GlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIleSerGlnGlnAla--- 337
QY 895 TACGATGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954
Db   :|||
QY 338 -----ThrSerProArgHisArgGly---GlyGluArg 347
QY 955 GGCAGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAACATTTGGT 1014
Db   :|||
QY 348 LeuAlaIleAsnLeuLeu----- 353
QY 1015 AGAAACAGATCCCTGACATCTACAACCTTCAAGCTGGTTCACTCAAACTGCCAACGAT 1074
Db   :|||
QY 354 ---SerGlnSerPro---ValTyrSerAsnGlnAsnGlyArgPhePheGluAlaCysPro 371
QY 1075 CTCAACCTTCTAATACCTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1134
Db   :|||
QY 372 GluAspPheSerGlnPheGlnAsnMetAspValAlaValSerAlaPheLysLeuAsnGln 391
QY 1135 AATGCATTTGTTTCGCTCACTACAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1194
Db   :|||
QY 392 GlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValValPheValThrAsp 411
QY 1195 GGACGGGCTCAGCTGCAAGTCTGTG-----CACAGCAACGGCAAC--- 1233
Db   :|||
QY 412 GlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSerGlnGlySerGln 431
QY 1234 -----AGAGTGTACGACGAGAGCTTCAAG----- 1260
Db   :|||
QY 432 SerGlyArgGlnAspArgArgGluGlnGluGluSerGluGluGluThrPheGlyGlu 451
QY 1261 -----GGTCAGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1290
Db   :|||
QY 452 PheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAlaProAlaGlyHis 471
QY 1291 GCCGTC-----GCTGGAAGTCCCAGAGCGAGAAGTTCGAATACGTTGGCATTTCAAG 1341
Db   :|||
QY 472 AlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAlaPheGlyLeuAsn 491
QY 1342 ACAGACTCAAGGCCGACATAGCAACCTCGCGGTGAAACCTCCGTCATAGATAACCTG 1401
Db   :|||
QY 492 AlaGlnAsnAsnGlnArgIlePhe---LeuAlaGlyLysLysAsnLeuValArgGlnMet 510
QY 1402 CCGAGAGAGTGTGTGCAAAATTCATATGGCTCCAAAGGAGGAGCAGCAAGCGAGCTTAAG 1461
Db   :|||
QY 511 AspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuValAspAsnIlePhe 530
QY 1462 AACAAACACCCCTCAAGTTCTTCGTTCCACCGCTCTCAGCAGTCTCCGAGG 1512
Db   :|||
QY 531 AsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArgGlnArg 547
```

RESULT 14

```
US-07-955-905A-22
; Sequence 22, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 566 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Theobroma cacao
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..566
; OTHER INFORMATION: /note="67 kD Precursor Protein"
; US-07-955-905A-22
```

```
Alignment Scores:
Pred. No.: 3,91e-06 Length: 566
Score: 156.00 Matches: 106
Percent Similarity: 34.82% Conservative: 74
Best Local Similarity: 20.50% Mismatches: 205
Query Match: 5.80% Indels: 132
DB: 1 Gaps: 22
```

US-10-728-323-3 (1-1524) x US-07-955-905A-22 (1-566)

```
QY 253 TCCTCTAGACATATCAAGAGCTCACACACAGGTCGTCAGTCCCAAGACCA 312
Db   :|||
QY 66 CysGluArgGluTyrLysGluGlnArgGln---GlnGluGluGluGlnArgGln 84
Db   :|||
QY 313 CCAAGACGCTCTCAAGGAGAAGACCAAGCCCAACAGCAA----- 351
Db   :|||
QY 85 TyrGlnGlnCysGlnGlyArgCysGlnGluGlnGlnGlnGlnArgGlnGlnGln 104
QY 352 -----CGAGATAGTCCACAGAAAGGTG 372
Db   :|||
QY 105 CysGlnArgLysCysTrpGluGlnTyrLysGluGlnArgGlyGluHisGluAsnTyr 124
QY 373 CACCGTTTC-----GATGAGGT-----GATCTCAT 399
Db   :|||
QY 125 HisAsnHisLysLysAsnArgSerGluGluGluGluGlnGlnArgAsnAsnProTyr 144
QY 400 GCAGTTTCCACCGGTGTTCTTTCTGGCTC---TACACGACGACGACACTGATGTTGT 456
Db   :|||
QY 145 TyrPheProLysArgArgSerPheGlnThrArgPheArgAspGluGluGlyAsnPheLys 164
QY 457 GCTGTTTCTTCTACTGACACCAAC-----AACGACAAACAGCTT 498
Db   :|||
QY 165 IleLeuGlnArgPheAlaGluAsnSerProProLeuLysGlyIleAsnAspTyrArgLeu 184
QY 499 GATCAGTTC-----CCGAGGAGATTCAATTTGGCTGGGAACACGAGCAAGAG--- 546
Db   :|||
QY 185 AlaMetPheGluAlaAsnProAsnThrPheIleLeuProHisCysAspAlaGluAla 204
QY 547 -----TTCTTAAGGTACCAAGCAACAAAGC 570
Db   :|||
QY 205 IleTyrPheValThrAsnGlyLysGlyThrIleThrPheValThrHisGluAsnLysGlu 224
QY 571 AGACAAGACGACGAGA-----AGCTTACCATAT---AGCCATACAGCCCG 615
Db   :|||
QY 225 SerTyrAsnValGlnArgGlyThrValValSerValProAlaGlySerThrValTyrVal 244
QY 616 CAAAGTCAACCTAGACAAGAAGCGGTGAATTTAGC-----CCTCAGGACAG 663
Db   :|||
QY 245 ValSerGlnAspAsnGlnGluLysLeuThrIleAlaValLeuAlaLeuProValAsnSer 264
```

QY 664 CACAGCCGACGAGCAGGACGAGCAAGAAACGAGAGCTGGAAACATCTTCAGC 723
Db 265 ProGlyTyrGluLeuPhePheProAlaGlyAsnAsnLysProGluSerTyrGly 284
QY 724 GGCTTCACGCGGAGTTCCTGGAACAAAGCCTTCAGAGTTACGAC- 774
Db 285 AlaPheSerTyrGluValLeuGluThrValPheAsnThrGlnArgGluLeuGluGlu 304
QY 775 ATAGTGCACAACTAAGAGGCGGACGAGAGTGAAGAAGAGGAGGAGCCATGTGACAGTG 834
Db 305 IleLeuGluGluGlnArgGlyGlnLysArgGlnGlnGlyGln- 318
QY 835 AGGCGAGCCCTCAGATCTTCGAGCCGATAGAAAGAGAGCTGCCGACGAGAGAGGAA 894
Db 319 GlnGlyMetPheArgLysAlaLysProGluGlnIleArgAlaIleSerGlnGlnAla- 337
QY 895 TACGATGAAGATGAATATGATACGATCAAGAGGATAGAAGGCGTGGCAGGGAAGCAGA 954
Db 338 -----ThrSerProArgHisArgGly- 347
QY 955 GGCAGGGGGATGATATTGAAGAGACGATCTGCACCGCAAGTGTATAAAGAACATTGGT 1014
Db 348 LeuAlaIleAsnLeu- 353
QY 1015 AGAACAGATCCCTGACATCTACACCCCTCAGCTGTTCTCACTCAAACTGCCAACGAT 1074
Db 354 ---SerGlnSerPro---ValTyrSerAsnGlnAsnGlyArgPhePheGluAlaCysPro 371
QY 1075 CTCACCTCTTAATACCTAGTGGCTGGACCTAGTCTGAATATGGAATCTCTACAGG 1134
Db 372 GluAspPheSerGlnPheGlnAsnMetAspValAlaValSerAlaPheLysLeuAsnGln 391
QY 1135 AATGCAATGTTTGTGCTACTACAAACCAACGACACACGATCATATATCGATTGAGG 1194
Db 392 GlyAlaIlePheValProHisTyrAsnSerLysAlaThrPheValPheValThrAsp 411
QY 1195 GGACGGCTCAGTCGAGTCTGTG- 1233
Db 412 GlyTyrGlyTyrAlaGlnMetAlaCysProHisLeuSerArgGlnSerGlnSerGln 431
QY 1234 -----AGAGTGTACGACGAGAGCTTCAAGAG- 1260
Db 432 SerGlyArgGlnAspArgGluGlnGluGluSerGluGluGluThrPheGlyGlu 451
QY 1261 -----GGTCACGTGTTGTGGTGGCCACAGAACTTC 1290
Db 452 PheGlnGlnValLysAlaProLeuSerProGlyAspValPheValAlaProAlaGlyHis 471
QY 1291 GCGCTC-----GCTGGAAAGTCCGACGAGCACTTCGAAATAGTGGCATTCAG 1341
Db 472 AlaValThrPhePheAlaSerLysAspGlnProLeuAsnAlaValAlaPheGlyLeuAsn 491
QY 1342 ACAGACTCAAGGCGCCAGATAGCAACCTCGCCGTCGAAACCTCCGTCATAGATACTG 1401
Db 492 AlaGlnAsnAsnGlnArgIlePhe- 1461
QY 1402 CCGGAGAGGTGGTGCATTCATATGCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1461
Db 511 AspSerGluAlaLysGluLeuSerPheGlyValProSerLysLeuValAspAsnIlePhe 530
QY 1462 AACACAAACCCCTCAAGTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGG 1512
Db 531 AsnAsnProAspGluSerTyrPheMetSerPheSerGlnGlnArg 547

RESULT 15

US-09-216-393B-81
; Sequence 81, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhaussen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B

; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 81
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
; ORGANISM: Toxoplasma gondii
US-09-216-393B-81

Alignment Scores:
Pred. No.: 4,07e-06 Length: 611
Score: 156.00 Matches: 92
Percent Similarity: 35.11% Conservative: 73
Best Local Similarity: 19.57% Mismatches: 169
Query Match: 5.80% Indels: 137
DB: 4 Gaps: 14

US-10-728-323-3 (1-1524) x US-09-216-393B-81 (1-611)

QY 29 AGTTCAGCGCTCAATGCGCAGACCTGACAAATCGCATTTGAATCAGAGCGGCTTACA 88
Db 216 SerSerProIleSerValThrAsnAspValTyrAlaLeuPheAspSerSerAlaSerPro 235
QY 89 TTGAGACTTGGAAACCCCAACACACGAGAGTTCGAATGCGCGCGGCTCTCTCGCT 148
Db 236 LeuHisAlaGlyLeuLeuSerSerLeuProGlyAlaValSerAlaSerGluArgLeu-Le 255
QY 149 TACTCTCCG-----CCGCAACGCCCTTCGTAGGCCCTTTCTACTCCAACTGCTC 196
Db 255 uThrAlaProAlaGluIleGlyProSerAlaSerAlaCysLeuSerValSerCysG1 275
QY 197 CCACGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTTGATATTCCCTGGTTGTC 256
Db 275 yProGlyGluMetSerProThrAlaAsp-----ThrTh 286
QY 257 CTAGACAC-----TATGAAGAGCTTCACACCAAGGTC 289
Db 286 r-ArgHisAspAlaGluGluArgGluArgAlaGluGluLysGluArgGlu 306
QY 290 GTCGATCTCAGTCCCAACACACACGAGCTTCGAAGGAGAGCAACCAAGCAAGCAGC 349
Db 306 rGlnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 326
QY 350 AACGAGATAGTCCACAGAAAGTGCACCTGTTTCATGAGGCTGATCTCATTTGCATTTCCCA 409
Db 326 lnglGluGluGluArgGluArgValGluGlu----- 338
QY 410 CCGGTGTTGTTTCTCGCTCTACAACGACCAACACACACACACACACACACACACAC 469
Db 338 ----- 338
QY 470 CTGACACCAACACACACACACACACCTTGATCAGTTCCCGCAGGATTCATTTGCTG 529
Db 339 -----GluLysAlaArgGlnArgGluGluGluGluGluGluGluGluGluGluGluGlu 352
QY 530 GGAACACGAGCAAGAGTCTTAAAGGTACCAACCAACCAACCAACCAACCAACCAACCAAG 589
Db 353 --ArgValGluGluGluLysAlaArgGlnArgGluGluGluGluGluGluGluGluGlu 371
QY 590 GCTTACCATATAGCCCATACGCCCAAGTTCAGCTAGCAAGAGAGCGGTGAATTTA 649
Db 372 -----ValGluGluGluLysAlaArgGlnArgGluGluGluGluGluGluGluGluGlu 387
QY 650 GCCTCGAGGACACACACCGCAGAGACGACGACGACGACGACGACGACGACGACGACGAC 709
Db 387 luArgArgValGluGluGluGluLysAlaArgGlnArgGluGluGluGluGluGluGlu 407
QY 710 GAAACATCTTCAGCGGCTTCAGCGGAGTTCCTGGAACCAAGCTTCCAGGTTGACGACA 769
Db 407 rgArgArgValGluGluGluLysAlaArgGlnArgGluGluGluGluGluGluGluGluGlu 427

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: August 24, 2005, 10:05:17 ; Search time 98.2593 Seconds
(without alignments)
12146.987 Million cell updates/sec

Title: US-10-728-323-3
Perfect score: 2691
Sequence: 1 CGGCAGCAACCGAGGAGGAGAA.....ctccgaggcgctggcgttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 3518262

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10728323/runat_23082005_124357_29323/app_query.fasta_1.4757
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAFSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10728323@cgn_1_1_221@runat_23082005_124357_29323
-NCPU=6 -ICPU=3 -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCKS=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubppa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubppa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	2673	99.3	507	15	US-10-245-871-96	Sequence 96, Appl
2	2673	99.3	507	15	US-10-253-286-96	Sequence 96, Appl
3	2673	99.3	510	14	US-10-228-806-6	Sequence 6, Appl
4	2673	99.3	510	15	US-10-100-303A-90	Sequence 90, Appl
5	2638	98.0	526	9	US-09-731-221-79	Sequence 79, Appl
6	2621	97.4	530	17	US-10-899-551-6	Sequence 6, Appl
7	1665	61.9	351	17	US-10-899-551-58	Sequence 58, Appl
8	1502	55.8	481	10	US-09-759-967-21	Sequence 21, Appl
9	1502	55.8	481	15	US-10-424-599-171702	Sequence 171702, A
10	1502	55.8	488	15	US-10-425-114-43865	Sequence 43865, A
11	1502	55.8	491	15	US-10-425-114-44047	Sequence 44047, A
12	1502	55.8	492	15	US-10-425-114-43836	Sequence 43836, A
13	1502	55.8	492	15	US-10-425-114-46294	Sequence 46294, A
14	1499.5	55.7	466	15	US-10-425-114-71959	Sequence 71959, A
15	1499.5	55.7	466	15	US-10-425-114-71973	Sequence 71973, A
16	1499.5	55.7	485	10	US-09-759-967-20	Sequence 20, Appl
17	1499.5	55.7	488	15	US-10-425-114-45645	Sequence 45645, A
18	1499.5	55.7	489	15	US-10-425-114-43983	Sequence 43983, A
19	1499.5	55.7	489	15	US-10-425-114-44676	Sequence 44676, A
20	1499.5	55.7	489	15	US-10-425-114-45612	Sequence 45612, A
21	1499.5	55.7	489	15	US-10-425-114-50117	Sequence 50117, A
22	1499.5	55.7	489	15	US-10-425-114-50207	Sequence 50207, A
23	1499.5	55.7	489	15	US-10-425-114-51412	Sequence 51412, A
24	1499.5	55.7	489	15	US-10-425-114-58429	Sequence 58429, A
25	1499.5	55.7	489	15	US-10-425-114-71902	Sequence 71902, A
26	1499.5	55.7	489	15	US-10-425-114-71909	Sequence 71909, A
27	1499.5	55.7	489	15	US-10-425-114-71957	Sequence 71957, A
28	1499.5	55.7	489	15	US-10-425-114-71958	Sequence 71958, A
29	1499.5	55.7	489	15	US-10-425-114-71967	Sequence 71967, A
30	1499.5	55.7	489	15	US-10-425-114-71995	Sequence 71995, A
31	1499.5	55.7	489	15	US-10-425-114-72287	Sequence 72287, A
32	1499.5	55.7	489	15	US-10-425-114-72288	Sequence 72288, A
33	1498	55.6	476	17	US-10-425-114-45639	Sequence 45639, A
34	1496	55.6	476	17	US-10-408-993-1	Sequence 1, Appl
35	1496	55.6	477	17	US-10-408-993-15	Sequence 15, Appl
36	1496	55.6	485	17	US-10-408-993-9	Sequence 9, Appl
37	1496	55.6	485	17	US-10-408-993-13	Sequence 13, Appl
38	1496	55.6	495	15	US-09-759-967-2	Sequence 2, Appl
39	1496	55.6	495	15	US-10-424-599-171705	Sequence 171705, A
40	1496	55.6	495	15	US-10-425-114-45626	Sequence 45626, A
41	1496	55.6	501	15	US-10-425-114-43904	Sequence 43904, A
42	1496	55.6	505	15	US-10-425-114-71930	Sequence 71930, A
43	1496	55.6	505	15	US-10-425-114-71962	Sequence 71962, A
44	1496	55.6	506	15	US-10-425-114-43901	Sequence 43901, A
45	1496	55.6	506	15	US-10-425-114-45624	Sequence 45624, A

ALIGNMENTS

RESULT 1

US-10-245-871-96
; Sequence 96, Application US/10245871
; Publication No. US20030235594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT

ORGANISM: Arachis hypogaea
US-10-245-871-96

Alignment Scores:
Pred. No.: 1,83e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 15 Gaps: 0

US-10-728-323-3 (1-1524) x US-10-245-871-96 (1-507)

QY	1	CGGAGCAACGGAGGAGAACGGTGCAGTTCAGCGCTCAATGCGCAGACCTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCATGAATCAGAGGCGGTTCATTAGACTTGGAAACCCCAACACGAGGATTC	120
DB	21	AsnArgileGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGGCGGCGTCCCTCTCTCGTTCAGTTCCTCGCGCGCAACGCCCTTCGTAGCCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTTG	240
DB	61	PheTy-SerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrllePheGlyLeu	80
QY	241	ATATTCCTGGTTGTCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG	300
DB	81	IlePheProGlyCysProArgHisTyrlleGluGluProHisThrGlnGlyArgSerGln	100
QY	301	TCCCAAGACCAACAGACCTCCAGGAGAGACCAAGCCCAACAGCAACGAGTACT	360
DB	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACCAAGGTGCACCGTTCGATGAGGGTGATCTCATTGCAGTTCACCGCGGTGTGCT	420
DB	121	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGGCTTACACGACCAACGACACTGATGTTGTTGTTCTCTTACTGACACAAC	480
DB	141	PheTrpLeuTyrlleAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
QY	481	AACAACGACACCGCTTCGATGATTCCTCCAGAGATTCATTTGCTGGGACACGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrlleGlnGlnSerArgGlnSerArgArgSerLeuProTyrlle	200
QY	601	AGCCCATACAGCCGCAAGTACCTAGCTAGCAAGAGAGCGTGAATTTAGCCCTCAGGA	660
DB	201	SerProTyrlleSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCCAGAGCAACGACGACGACAGAGAGAAACCAAGGTGGAACATCTTC	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGlnGluAsnGluGlyAsnIlePhe	240
QY	721	AGCGGCTTACCGCGGAGTTCCTGGAAACAGCCTTCAGGTTCCAGCAGACAGATAGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAAACCTAAGAGCGGAGACCGAGTAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280
QY	841	GGCCTCAGATCTTGACCCAGATAGAAAGAGACGTCCGACCAAGAGAGGAATACGAT	900
DB	281	GlyLeuArgileLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrlleAsp	300

QY	901	GAAGATGAATATGAATACGATGAAGAGGATGAAGCGGTGGCAGGGGAGAGAGCGAGG	960
DB	301	GluAspGluTyrlleGluTyrlleAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGTTATGAAGAGACGATCTGCACCCGCAAGTGTCTAAAGAAACATTGGTAGAAC	1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
QY	1021	AGATCCCTTGACATCTACAACCCCTCAAGCTGTTTCACCTCAAACTGCCAACCATCTCAAC	1080
DB	341	ArgSerProAspIleTyrlleAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAATATCTTAGTGGCTTCGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA	1140
DB	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrlleGlyAsnLeuTyrlleArgAsnAla	380
QY	1141	TTGTTTTCGCTCACTACAACACCAACCGCACACATCATATATCGATTGAGGGGACGG	1200
DB	381	LeuPheValAlaHisTyrlleAsnThrAsnAlaHisSerIleIleTyrlleArgLeuArgGlyArg	400
QY	1201	GCTCAGCTGCAAGTCTGTCGACAGCAACCGCAACAGAGTGTACGAGGAGCTTCAAGAG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrlleAspGluGluLeuGlnGlu	420
QY	1261	GCTCAGCTGCTTGTGTCGCCACAGAACTTCGCCGTCGCTGGAAAGTCCAGAGCGAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTGCAATAGCTGGCATTCAAGACACTCAAGGCCCGCCAGCATAGCCAACTCCCGCGTGA	1380
DB	441	PheGluTyrlleValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCCGTCATAGATAAACCCTCGCGAGGAGGTGTTGCAAAATTCATATGGCCTCCAAAGG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrlleGlyLeuGlnArg	480
QY	1441	GAGCAGGCAAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGGCT	1521
DB	501	GlnSerProArgAlaValAla	507

RESULT 2

US-10-253-286-96
; Sequence 96, Application US/10253286
; Publication No. US20040058861A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-96

Alignment Scores:
Pred. No.: 1,83e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0

DB: 15 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-253-286-96 (1-507)
QY 1 CGGCAGCAACCGGAGGAGAACCGCTGCGCAGTTCACAGCGCCCTCAATGCGCAGACCTGAC 60
DB 1 ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProAsp 20
QY 61 AATCGCATTTGAATCAGAGGGGGTTCATTGAGACTTGGAAACCCCAACAACACGAGATTC 120
DB 21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAenGlnGluPhe 40
QY 121 GAATGCGCGCGCGCTCTCTCGCTTACTCGCTTACTCGCGCAGCAACCGGATCTTGGCGCT 180
DB 41 GluCysAlaGlyAlaAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 60
QY 181 TTCTACTCCCAATGCTCCCGAGGAGATCTTCATCAGCAAGCAAGGAGGATCTTGGCGTTG 240
DB 61 PheTyrSerAenAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY 241 ATATTCCCTGTTTCTCTAGACACTATGAAGAGCTCACACAAAGTGTGCTGATCTCAG 300
DB 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 100
QY 301 TCCCAAGACACCAAGCGTCTCCAAGGAGAACCAAGCCACAGCAACGAGATAGT 360
DB 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY 361 CACCAGAGGTGACCGTTTCGATGAGGTGATCTCATTTGAGTTCACCGGTTGCT 420
DB 121 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY 421 TTCTGGCTCTACAGCACCGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
DB 141 PheTrpLeuTyrAenAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY 481 AACCAACGACACCGCTTGATCAGTCCCGAGGATTCATTTGGCTGGCAACCGGAG 540
DB 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAenLeuAlaGlyAenThrGlu 180
QY 541 CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
DB 181 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 200
QY 601 AGCCCATACAGCCGCAAGTACGCTTAGACAAGAGCGGTGAATTTAGCCCTCGAGGA 660
DB 201 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 220
QY 661 CAGCAGACCCGACAGACGAGCAGGACAGAGCAAGAAACGAGGTGGAACATCTTC 720
DB 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAenIlePhe 240
QY 721 AGCGCTTCCGCGGAGTTCCTCGAAACAGCCCTCCAGGTTTACAGCAGACAGATAGTG 780
DB 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY 781 CAAACCTTAAGCGCAGACCGAGTAGTGAAGAGAGGAGGCCATTGTACAGTGAAGGGA 840
DB 261 GlnAenLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
QY 841 GGCTTCAGATCTTGACCCCGACAGATAGAAAGAGAGCTGCCAGCAGAGAGGATACGAT 900
DB 281 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 300
QY 901 GAAGATGAATATGAATACGATGAAGAGCATAGAAGCGGTGCGAGGGAAGCAGAGGAGG 960
DB 301 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 320
QY 961 GGGAAATGGTATTGAAGAGACGATTCGACCCGCAAGTCTAAAAGAAACATTGGTAGAAC 1020
DB 321 GlyAenGlyIleGluGluThrIleCysThrAlaSerAlaIysLysAenIleGlyArgAsn 340
QY 1021 AGATCCCTGACATCTACAACCCCTCAAGCTGGTTCATCTCAAAACTGCCAAGCATCTCAAC 1080

DB 341 ArgSerProAspIleTyrAenProGlnAlaGlySerLeuLysThrAlaAenAspLeuAen 360
QY 1081 CTTCTAATATCTAGTGGCTTGGACCTAGTGCCTGAATATGAAATCTCTACAGAAATGCA 1140
DB 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAenLeuTyrArgAenAla 380
QY 1141 TTGTTTGTCTCCTCACTAACACACACGACGACAGCATCATATATCATTTGAGGGGACGG 1200
DB 381 LeuPheValAlaHisTyrAenThrAenAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY 1201 GCTCAGTGAAGTCTGTCGACACGACGACGACGAGTGTACGACGAGGAGCTTCAAGAG 1260
DB 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
QY 1261 GGTCACTGCTTGTGTGTCACAGAACTTCGCGCTCGCTCGAAAGTCCACAGAGCGAAC 1320
DB 421 GlyHisValLeuValValProGlnAenPheAlaValAlaGlySerGlnSerGluAen 440
QY 1321 TTGCAATATCTGTCATTAAGACAGACTCAAGGCCGACGATAGCCAACTTCGCGGTGAA 1380
DB 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAenLeuAlaGlyGlu 460
QY 1381 AACTCCGTCACTAGATAACCTGCGGAGGAGTGTGTCGAAATTCATATGCGCTCCAAAGG 1440
DB 461 AsnSerValIleAspAenLeuProGluGluValAlaAenSerTyrGlyLeuGlnArg 480
QY 1441 GACGAGCAGGACGAGCTTAAGAAACAACACCCCTTCAAGTCTTCGTTCCACCGTCTCAG 1500
DB 481 GluGlnAlaArgGlnLeuLysAenAenProPheLysPhePheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
DB 501 GlnSerProArgAlaValAla 507
RESULT 3
US-10-228-806-6
; Sequence 6, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-6
Alignment Scores:
Pred. No.: 1,83e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 14 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-228-806-6 (1-510)
QY 1 CGGCAGCAACCGGAGGAGAACCGCTGCGCAGTTCACAGCGCTCAATGCGCAGACCTGAC 60
DB 4 ArgGlnGlnProGluGluAenAlaCysGlnPheGlnArgLeuAenAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGGGGGTTCATTGAGACTTGGAAACCCCAACAACACGAGATTC 120
DB 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAenAenGlnGluPhe 43
QY 121 GAATGCGCGCGGCTGCGCTCTCTCGCTTACTCGCTTACTCGCGGCAACCGCCCTTCGTAGGCCT 180

Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAGGGATCTTTGGGTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACACAGGTCTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACAAGACTCTCCAAGGAGAGACAAAGCCAAAGCCAAACAGCATAGT 360
Db 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTTCCACCGGTGTGCT 420
Db 124 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAAACGACACGACTGATGTTGTTGCTGTTTCTTCTACTGACACCAAC 480
Db 144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACAAACCAACGACTTGATTCAGTTCCCCAGAGATTCAATTTGCTGGGAACACGGAG 540
Db 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCTTAAGTACCAGCAACAAGCAGACAAAGCAGACGAAGAGCTTACCATAT 600
Db 184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY 601 AGCCCATACAGCCCGCAAGTCTAGCCTAGACAAGAGAGCGTCAATTTAGCCCTCGAGGA 660
Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223
QY 661 CAGCACAGCCGACAGAGAACGAGCAGCAAGAAGAGAAACGAAGGTGGAAACATCTTC 720
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 243
QY 721 AGCGGCTTACCGCGGNGTTCTCGAACAGCCTTCCAGGTTCCACGACAGACAGATAGT 780
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAAACCTAAGAGCGCAGACCCAGAGAGTGAAGAAGAGAGGCCATTTGACAGTGGGGA 840
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
QY 841 GGCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTCGCCGACCAAGAGGAATACGAT 900
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
QY 901 GAAGATCAATATGAATACGATGAAGGATAGAGCGGTGGCAGGGGAGCAGGCGAG 960
Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 323
QY 961 GGGAAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAAAGAACATTTGTAGAAAC 1020
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIysIysAsnIleGlyIArgAsn 343
QY 1021 AGATCCCTGACATCTACAAACCTCAAGCTGGTTCTCACTCAAACTGCCAAACGATCTCAAC 1080
Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuIysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATACTAGTGGCTTGACCTAGCTGAATATGAAATCTCTACAGGAATGCA 1140
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
QY 1141 TTGTTTCTCCTCACTACACACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY 1201 GCTCAGTGAAGTCTGTGGAACGAAACGAGGTGTACACGAGGAGCTTCAAGAG 1260

Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY 1261 GGTACGCTGCTGTTGGTGCACAGAACTTCGCGCTCGCTGGAAAGTCCACAGACGAGAAC 1320
Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyIysSerGlnSerGluAsn 443
QY 1321 TTCGAATACGTGGCAATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
Db 444 PheGluTyrValAlaPheIysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCGCTCATAGATAACCTGCCGAGGAGGTGGTTCGAAATTCATATGCGCTCCAAAG 1440
Db 464 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuGlnArg 483
QY 1441 GAGCAGGCAAGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
Db 484 GluGlnAlaArgGlnLeuIysAsnAsnProPheIysPhePheValProProSerGln 503
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 504 GlnSerProArgAlaValAla 510
RESULT 4
US-10-100-303A-90
; Sequence 90, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-90
Alignment Scores:
Pred. No.: 1.83e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.33% Indels: 0
DB: 15 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-100-303A-90 (1-510)
QY 1 CGGCAGCAACCGAGGAGAACCGTGCAGTTCAGCGCCTCAATCGCGAGAGACCTGAC 60
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGCGGTTTACATTGAGACTTGGAAACCCCAACACAGGATTC 120
Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTyrAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGGCGTGCCTCTCTCGCTTAGCTTCCTCCGCGCAACGCCCTTGTAGGCCT 180
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 63
QY 181 TTCTACTCCAAATGCTCCCAAGGAGATCTTCATCCAGCAAGGAGGGATCTTTGGGTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGTTGCTCTAGACACTATGAAGAGCCTCACACACAGGTCTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
QY 301 TCCCAAGACCAACAAGACTCTCCAAGGAGAGACAAAGCCAAAGCCAAACAGCATAGT 360

104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
Db
361 CACGAGAGGTGACCGTTCATGACAGGTGATCTCATTCAGTGTCCACCGGTGTGCT 420
Qy
124 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
Db
421 TTCTGGCTCTACACGACGACGACACTGATGTTCTGCTGTTCTTCTTACTGACCAAC 480
Qy
144 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
Db
481 AACACGACACCAAGCTTGATTCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAAACACCGAG 540
Qy
164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
Db
541 CAAGAGTCTTAAGGTACCGACAAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Qy
184 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 203
Db
601 AGCCCATACAGCCGCAAGTGCAGCTAGACAGAGCGTGAATTTAGCCCTCGAGGA 660
Qy
204 SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly 223
Db
661 CAGCACAGCCGACAGAACGAGCAGGACAAAGAAAGAAAGAGGTGAAACATCTTC 720
Qy
224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
Db
721 AGCGGCTTACGCGGAGTCTTCGAAACAAAGCTTCAGGTTGACGACAGACAGATAGTG 780
Qy
244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
Db
781 CAAACCTAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGA 840
Qy
264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
Db
841 GGCCTCAGAATCTTTAGCCAGATAGAAAGAGACGTCGCCACGAAGAAGAGGAATACGAT 900
Qy
284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
Db
901 GAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy
304 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgLysArg 323
Db
961 GGGAACTGATTCAGAGAGAGTCTGACCGCAAGTCTAAAGACATTCGTAGAAC 1020
Qy
324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyLysAsn 343
Db
1021 AGATCCCTCAGCATCTACAAACCTCTCAAGCTGGTTCCTCAAAACTGCCAACGATCTCAAC 1080
Qy
344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
Db
1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGATGATGATGATGATGATGATGATGATGAT 1140
Qy
364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGlyTyrGlyAsnLeuTyrArgAsnAla 383
Db
1141 TTCTTTGCTCCTCCTACACACCAAGCAGCAGCATCATATATCATATGATGAGGACCG 1200
Qy
384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
Db
1201 GCTCAGCTGCAAGTCTGTGGACAGCAACCGCAACAGAGTGACACAGAGAGCTTCAAGAG 1260
Qy
404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluLeuGlnGlu 423
Db
1261 GGTCACTGCTTGTGTGTCACAGAACTTCGCCGTCCGCTGGAAGTCCACAGCGAGAAC 1320
Qy
424 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
Db
1321 TTGGAATACGTGCATTCAGACAGACTCAAGCCCGCAGCATACCCACCTCGCCGTGAA 1380
Qy
444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
Db
1381 AACTCCGTCTATAGTAACTTCGCGGAGGAGTGGTTCGAATTCATATGSCCTCCAAAGG 1440
Qy
464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
Db

1441 GAGCAGCGAAGGACGCTTAAGAAACAACACCCCTTCAAGTTCCTGTTCCACCGTCTCAG 1500
Qy
484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 503
Db
1501 CAGTCTCCGAGGGCTGTGGCT 1521
Qy
504 GlnSerProArgAlaValAla 510
Db
RESULT 5
US-09-731-221-79
; Sequence 79, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-79
Alignment Scores:
Pred. No.: 1,94e-218 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 5
Query Match: 98.03% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-3 (1-1524) x US-09-731-221-79 (1-526)

1 CGCAGACAAACCGGAGGAGAAACGCGTCCAGTTCAGCGCTCAATGCGAGAGACCTGAC 60
Qy
5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
Db
61 AATCGCATTTGAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACCAAGAGATTC 120
Qy
25 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
Db
121 GAATGCGCGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAGACGCCCTTCGTAGGCCT 180
Qy
45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 64
Db
181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGAAAGGAGGATCTTTGGGTG 240
Qy
65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
Db
241 ATATTCCCTGGTGTCTTAGACACTATGAGAGCCTCACACAAAGTCTCGATCTCAG 300
Qy
85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
Db
301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGAACCAAGCAACAGCAAGATAGT 360
Qy
105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
Db
361 CACCAAGCTGCACCGTTCGATGAGGCTGATCTCATTCAGATTCGCCCGGTGTGCT 420
Qy
125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
Db
421 TTCTGCTCTACACGACCAACGACACTGATGTTGTTGTTGTTGTTTCTTACTGACCAAC 480
Qy
145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
Db
481 AACACGACAAACAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGCTGGGAACACGGAG 540
Qy

Db	165	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAenThrGlu	184
QY	541	CAAGAGTCTTAAGTACCAAGCAACAAAGCAGACGAAGAGCTTACCATAT	600
Db	185	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	204
QY	601	AGCCCATACAGCCCGCAAGCTAGCCTTAGACAAGAAGCGTGAAATTTAGCCCTCAGGA	660
Db	205	SerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArgGly	224
QY	661	CAGCAGCCGCGAGAACGAGCAGCAGAAGAAGAAAGAAACGAAGGTGGAACATCTTC	720
Db	225	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAenIlePhe	244
QY	721	AGCGGCTTCAGCCGCGAGTCTCTGGAACAAGCCCTCCAGGTTGACACACAGACATAGT	780
Db	245	SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal	264
QY	781	CAAAACCTTAAGAGCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGGGA	840
Db	265	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	284
QY	841	GGCCTCAGAACTTGTAGCCGAGATGAAGAAGACGTCGCCGACGAAGAAGAGGAATACGAT	900
Db	285	GlyLeuArgAlaLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	304
QY	901	GAAGATCAATATCAATACGATGAAGAGGATAGAAGCGCTGGCAGGGGAACGACGACGG	960
Db	305	GluAspGluTyrAlaTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	324
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGCTAAAAGAACATTTGGTAGAAAC	1020
Db	325	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleYsAsnIleGlyArgAsn	344
QY	1021	AGATCCCTGACATCTACAACCTCAAGCTGGTTCTACTCAAACTGCCAACGATCTCAAC	1080
Db	345	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	364
QY	1081	CTTCTAATACTTAAAGTGGCTTGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA	1140
Db	365	LeuLeuIleuArgTyrPheGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	384
QY	1141	TTGTTTTCGCTCACTACAACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
Db	385	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	404
QY	1201	GCTCAGTGCAGTCTGTGACGACACGCGCAACAGAGTGTAACGAGGAGCTTCAAGAG	1260
Db	405	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	424
QY	1261	GGTCAGTGTCTGTGTGCCACAGAACTTCGCCGTCGCTGGAAGTCCCGAGCGAGAAC	1320
Db	425	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	444
QY	1321	TTCGAATACGTGGCATTTCAAGACAGACTCAAGGCCCGCAGCATAGCCAACTCTCGCGGTGAA	1380
Db	445	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	464
QY	1381	AACCTCGTCATAGATAACCTGCCGAGGAGGTGGTTGCCAAATTCATATATGGCTCCAAAGG	1440
Db	465	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	484
QY	1441	GAGCAGCGAAGGAGCTTAAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGCTCTCAG	1500
Db	485	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln	504
QY	1501	CAGTCTCCGAGGGCTGTGGCT	1521
Db	505	GlnSerProArgAlaValAla	511
RESULT 6			
US-10-899-551-6			

; Sequence 6, Application US/10899551			
; Publication NO. US20050063994A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael J.			
; APPLICANT: Burks, A. Wesley			
; APPLICANT: Sampson, Hugh A.			
; APPLICANT: Howard, Sosin B.			
; APPLICANT: Bottomly, Kim H.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy			
; FILE REFERENCE: 2002834-0233			
; CURRENT APPLICATION NUMBER: US/10/899,551			
; CURRENT FILING DATE: 2004-07-26			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 6			
; LENGTH: 530			
; TYPE: PRT			
; ORGANISM: species Arachis hypogea			
US-10-899-551-6			
Alignment Scores:			
Pred. No.: 5,7e-217 Length: 530			
Score: 2621.00 Matches: 500			
Percent Similarity: 98.62% Conservatives: 0			
Best Local Similarity: 98.62% Mismatches: 7			
Query Match: 97.40% Indels: 0			
DB: 17 Gaps: 0			
US-10-728-323-3 (1-1524) x US-10-899-551-6 (1-530)			
QY	1	CGCGACCAACCGAGAGAGAACCGTGCACCGCTTCCAGCGCTCAATGCGCAGAGACCTGAC	60
Db	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	ATTCGCTTGAATCAGAGCGGCTTACATGAGACTTGAGATCTCGAACCCCAACACGAGGTTTC	120
Db	44	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAAGCCCTTCGTAGGCCT	180
Db	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro	83
QY	181	TTCTACTCCAAATGCTCCCGAGAGATCTTCAATCCAGGAAGAGGAGATCTTTGGGTG	240
Db	84	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	103
QY	241	ATATTCCCTGGTGTCTTAGACACTATGAAGAGCTCACACAAAGTCTCGATCTCAG	300
Db	104	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln	123
QY	301	TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGACCAAGCAACAGCAACGAGATAGT	360
Db	124	SerGlnArgProIleArgLeuGlnGlyLeuAspGlnSerGlnGlnGlnArgAspSer	143
QY	361	CACCAAGAGTGCACCGTTTCGATGAGGTATCTATTGCGATGTTCCCAACCGGTGTGCT	420
Db	144	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
QY	421	TTCTCGCTCTACAACGACCCAGCATGTATGTTGTGTTGTTCTTCTTACTCACACCAAC	480
Db	164	PheTrpLeuTyrAsnAspHisAspThrAspValAlaIleValSerLeuThrAspThrAsn	183
QY	481	AAACAACGACCAACGAGTTGATCAGTTTCCCGAGAGATTCAATTTGGTGGGAACGCGAG	540
Db	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
QY	541	CAAGAGTCTTAAAGGTACCAAGCAACAAAGCAGACAAAAGCAGACGAAGAAGCTTACCATAT	600
Db	204	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	223
QY	601	AGCCCATACAGCCCGCAAGTCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
Db	224	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	243


```
QY 661 CAGCAGCCGCGCAGACGAGCAGGACAGAGACAGAAACGAGAGTGGAAACATCTTC 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 263
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAACAAGCCCTCCAGGTTGACGACAGACAGATAGTG 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAACACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
Db 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGlyAlaIleValThrValArgGly 303
QY 841 GGCTCAGATCTTGACCCAGATAGAAAGAGAGCTGCCAGAGAGAGGAGATACGAT 900
Db 304 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluIuTyArg 323
QY 901 GAAGATCAATATGAATACGATGAAGAGCAGATAGAGGCTGCGAGGGGAACGAGGGCAGG 960
Db 324 GluAspGluTyArgGluTyArgGluGluAspArgArgGlyArgGlySerArgLysArg 343
QY 961 GGAATCGTATTGAAGAGAGATCTGCACCGCAAGTGTCTAAAGAAACATTGGTAGAAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 363
QY 1021 AGATCCCTGCATCTACACCCCTCAAGCTGGTTCACCTCAAACTGCCAAGCATCTCAAC 1080
Db 364 ArgSerProAspIleTyArgAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 383
QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGAATATGAAATCTCTACAGGAATGCA 1140
Db 384 LeuLeuIleuArgTrpLeuGlyProSerAlaGluTyArgLysAsnLeuTyArgAsnAla 403
QY 1141 TTGTTTTCGCTCACTACAACACCAACGACACAGCATCATATATCGATTGAGGGACGG 1200
Db 404 LeuPheValAlaHisTyArgAsnThrAsnAlaHisSerIleIleTyArgLeuArgGlyArg 423
QY 1201 GCTCAGTGAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGG 1260
Db 424 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyArgGluGluLeuGlnGlu 443
QY 1261 GGTCACTGTGTGTGGTGCACAGAACTTCGCGTGTGCTGGAAGTCCACAGAGCGAAGC 1320
Db 444 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 463
QY 1321 TTCGAATACGTGGCATTCAAGACAGACTCAAGGCCACAGCATAGCAACCTCGCGGTGAA 1380
Db 464 PheGluTyValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
QY 1381 AACTCGTCTATAGATAACCTGCGGAGGAGGTGTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 484 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyArgLysLeuGlnArg 503
QY 1441 GAGCAGCAAGGCGAGCTTAAGAACAAACCCCTTCAAGTTCTTCTCCACCGTCTCAG 1500
Db 504 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 523
QY 1501 CAGTCTCCGAGGGCTGTGGCT 1521
Db 524 GlnSerProArgAlaValAla 530
```

RESULT 7

```
US-10-899-551-58
; Sequence 58, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
```

```
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 351
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-58
```

Alignment Scores:

Pred. No.:	1,53e-134	Length:	351
Score:	1665.00	Matches:	315
Percent Similarity:	97.83%	Conservative:	0
Best Local Similarity:	97.83%	Mismatches:	7
Query Match:	61.87%	Indels:	0
DB:	17	Gaps:	0

US-10-728-323-3 (1-1524) x US-10-899-551-58 (1-351)

```
QY 1 CGGCAGCAACCGGAGGAGAAACGCGTCCAGTTCACGGCTCAATGCGCAGAGACTTGAC 60
Db 15 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 34
QY 61 AATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTGGAAACCCCAACACACGAGGATTC 120
Db 35 AsnArgIleGluSerGluGlyGlyIleGluThrTrpAsnProAsnAsnGlnGluPhe 54
QY 121 GAATCGCGCGGCTCCGCTCTCTCGCTAGTCTCCGCGCAACGCCCTTCGTAGGCTT 180
Db 55 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 74
QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGAGAGGAGATCTTTGGGTTG 240
Db 75 PheTySerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyPheGlyLeu 94
QY 241 ATATTCCTCGTGTCTAGACACTATGAGAGCTTACACAGAGCTTCCATCGATCTCAG 300
Db 95 IlePheProGlyCysProSerThrTyArgGluProAlaGlnGlnGlyArgArgSerGln 114
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAAGAGAGAGCAAAAGCCAAACAGCAACGAGATAGT 360
Db 115 SerGlnArgProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 134
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGTGTGCT 420
Db 135 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 154
QY 421 TTCTGCTCTACACGACCAACACGACTGATGTGTGCTGTTCTTCTTACTGACACCAAC 480
Db 155 PheTrpLeuTyArgAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 174
QY 481 AACCAACGACCAACCGCTTCATCATGTTCCCGAGAGATTCAATTTGGTGGGAACACGAG 540
Db 175 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 194
QY 541 CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 195 GlnGluPheLeuArgTyArgGlnGlnSerArgGlnSerArgArgSerLeuProTyArg 214
QY 601 AGCCCATACGCGCGCAAGTCAAGCTTACAGAGAGAGCGTGAATTTAGCCCTCGAGGA 660
Db 215 SerProTySerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 234
QY 661 CAGCAGCGCGGAGAGAACGAGCAGGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 235 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGluGlu 254
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCACAGGTTTGACGACAGACAGATAGTG 780
Db 255 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 274
QY 781 CAAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACAGTGAGGGGA 840
```


RESULT 9
US-10-424-599-171702
; Sequence 171702, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yihua
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)8
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 171702
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126061C.1.pep
US-10-424-599-171702

Alignment Scores:
Pred. No.: 1.99e-120 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 15 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-424-599-171702 (1-481)

QY 1 CGGCAGCAACCGGAGGAGACGGCTGCAGTTCAGCGCCTCAATGGCGAGACCTGAC 60
DB 23 ArgGluGlnProGlnGlnGlnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42

QY 61 AATCGCATTAATGATGAGAGCGGTTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62

QY 121 GAATGGCGCGGCTCGCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCT 180
DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82

QY 181 TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGGTTG 240
DB 83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102

QY 241 ATATTCCCTGGTTGCTAGACACTATGAAGAGCCTCACACAAAGGTGCTCGATCTCAG 300
DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119

QY 301 TCCCAAGACACCAAGACGCTCTCCAGGAGAGACCAAGGCCACAGCAACAGATAGT 360
DB 120 SerSerArgPro-----GlnAspArg 126

QY 361 CACCAAGAGTGACCGTTTCGATGAGGATGATCTCATTTCAGTTCACCGGTGTGCT 420
DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146

QY 421 TTCTGGCTCTACACGACACGACACTGATGTTGTGCTGTTCTCTTACTGACACCAAC 480
DB 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166

QY 481 AACACGACCAACGCTTGATGATGTTCCCGAGGAGATTCAAATTTGGCTGGGAACAGGAG 540
DB 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186

QY 541 CAAGAGTCTTAAGGTACCAAGCAACAAAGCAGACGAGCAGACGAGAGCTTACCATAT 600
DB 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199

QY 601 AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
DB 200 -----Gly 200

QY 661 CAGCACAGCGCAGAGAACGAGCAGGACGAGAACGAGAACGAGAACGAGAACGAGAACGAG 720
DB 201 ThrGlnSerGlnLysGlyArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220

QY 721 AGCGGCTTCCAGCGCGGAGTTCTTCGGAACCAAGCCTTCAGGTTTACGACACAGACAGTAGTG 780
DB 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239

QY 781 CAAAACCTTAAGAGCGCAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 240 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyAlaIleValThrValLysGly 259

QY 841 GGCTCTCAGAACTTTGAGCCCA-----GATAGAAAGAGAGCGTCCCGACGAGAGAGAG 891
DB 260 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlu 279

QY 892 GAATACGATCAAGATCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 951
DB 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294

QY 952 AGAGCGAGCGGGAATGTTTGAAGAGAGAGATCTGCACCGCAAGTGTCTAAAGAGACATT 1011
DB 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311

QY 1012 GGTAGAAAACAGATCCCTGCATCTACAACTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCA 1071
DB 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331

QY 1072 GATCTCAACCTTCTAATACCTTAGTGGCTTGGACCTAGTGTGATGATGATGATGATGATG 1131
DB 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351

QY 1132 AGAATGCAATGTTTTCGCTCACTACAACCAACGACACACAGCATCATATATATCATGATTG 1191
DB 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371

QY 1192 AGGGGAGCGGCTCAGCTGCAAGTCTGGGACAGCAAGCGCAACAGAGTGTACGACGAGGAG 1251
DB 372 AsnGlyArgAlaLeuValGlnValValAsnGlyGluArgValPheAspGlyGlu 391

QY 1252 CTTCAAGAGGCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTCGCTCGAAAGTCCAG 1311
DB 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411

QY 1312 AGCGAAGCTTCGAATACGTGGCATTCAGACAGACACTCAAGGCCAGCATAGCCAACTTC 1371
DB 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431

QY 1372 GCGGTGAAAACCTCGCTCATAGTACCTGCGGAGAGGTGTTGCCAAATTCATATGGC 1431
DB 432 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451

QY 1432 CTCCAAGGAGGAGCAGGCAAGGAGCTTAAGAACCAACCCCTTCAAGTTCTTCGTTCCTCA 1491
DB 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471

QY 1492 CGGTCTCAGCAGTCTCCGAGGCGTGTGGCT 1521
DB 472 ProLysGluSerGlnArgArgValValAla 481

RESULT 10
US-10-425-114-43865
; Sequence 43865, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.

```
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 43865
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700645818_FLI.pep
US-10-425-114-43865

Alignment Scores:
Pred. No.: 2e-120 Length: 488
Score: 1502.00 Matches: 297
Percent Similarity: 71.57% Conservative: 68
Best Local Similarity: 58.24% Mismatches: 91
Query Match: 55.82% Indels: 54
DB: 15 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-425-114-43865 (1-488)
QY 1 CGGCAGCAACGGGAGGAGACGGTCCAGTTCAGCGCTCAATGCGCAGAGACCTGAC 60
DB 30 ArgGluGlnProGlnAsnGluCysGlnLeuArgLeuAsnAlaLeuLysProAsp 49
QY 61 AATCGCATTAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
DB 50 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 69
QY 121 GAATGCGCGGCTGCGCTCTCTCGCTAGTCTCGCGCGCAACGCGCTTCGTAGCCT 180
DB 70 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 89
QY 181 TTCTACTCCAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGTTG 240
DB 90 SerTyrTrpAsnAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet 109
QY 241 ATATTCCTGTTGTTCTAGACACTATGAGAGCCTCACACAAAGGTGTCGATCTCAG 300
DB 110 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 126
QY 301 TCCCAAGACCAACAGACGCTCTCCAGGAGGAAGACCAAGCCACAGCAACGAGATAGT 360
DB 127 SerSerArgPro-----GlnAspArg 133
QY 361 CACCAGAGGTGCACCGTTTCGATAGGGTGATCTCATTGCAGTCCCAACCGGTGTGCT 420
DB 134 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 153
QY 421 TTCTGGCTCTACACACACACGACACTGATGTTGTTGCTTCTTCTTACTGACACCAAC 480
DB 154 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 173
QY 481 AACACGACACACAGCTTGATCAGTTCCCGCAGGAGATTCAAATTTGGCTGGGAACAGGAG 540
DB 174 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 193
QY 541 CAAGAGTCTTAGAGTACCAGCAACAAAGCAGACAGACGAGCAGACGAGAGCTTACCATAT 600
DB 194 GlnGluPheGluGlnTyrGlnProGlnLysGlnGly----- 206
QY 601 AGCCCATACACCGCGCAAAAGTCAGCTTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 207 -----Gly 207
QY 661 CAGCAGCGCGCAGACGAGGAGGAGCAAGAGAAACGAAAGTGGAAACATCTTC 720
DB -----
```

```
DB 208 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAsnGluGlyGlySerIleLeu 227
QY 721 AGCGCTTTCACGCCGAGTCTCTGAAACAAGCTTCCAGGTTGACAGACAGATAGTG 780
DB 228 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 246
QY 781 CAAAACCTTAAGAGGCGAGACGAGAGTGAAGAGAGGAGGCCATTGTGCACAGTGGGGA 840
DB 247 ArgLysLeuGlnGlyGluAsnGluGluGluLysGlyAlaIleValThrValLysGly 266
QY 841 GGCCTCAGAATCTTCAGCCCA-----GATGAAAGAGACGTCGCCGCAAGAGAG 891
DB 267 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGlu 286
QY 892 GAATACGATGAAGATCAATATGAAATACGATGAAGAGGATGAAGAGCGTGGCAGGGAAGC 951
DB 287 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 301
QY 952 AGAGCAGGGGGAATGTTATTCAGAGAGCAGTCTCCACCGCAAGTCTGCTAAAAGACATT 1011
DB 302 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 318
QY 1012 GGTAGAAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCTACTCAAAACTGCCAAC 1071
DB 319 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 338
QY 1072 GATCTCAACCTCTTAATATCTAGTGGCTTGACCTAGTGTGAATATGAAATCTCTAC 1131
DB 339 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 358
QY 1132 AGGAATGCAATTTGTTGCTCCTACACCAACGACGACAGCATCATATATCGATTG 1191
DB 359 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 378
QY 1192 AGGGCAGGGCTCAGTCGCAAGTCTGTGACACCAACGCAAGAGTGTACACGAGAG 1251
DB 379 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 398
QY 1252 CTTCAAGAGGGTCAAGTGTGTCGACACAGAACTTCGCGCTCGCTGGAAAGTCCAG 1311
DB 399 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 418
QY 1312 ACCGAACTTCGAATAGTGCATTCAGACAGACTCAAGCCCGCAGCATAGCCACCTC 1371
DB 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
QY 1372 GCGGTGAAAACCTCCGTCATAGATAACCTCGCGGAGGAGTGGTTGCATAATTCATATGC 1431
DB 439 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
QY 1432 CTCCAAAGGGAGCAGGAGGAGGAGTGAAGAAACAACAACCCCTTCAAGTTCTCGTTCCA 1491
DB 459 LeuArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 478
QY 1492 CCGTCTCAGCAGTCTCCGAGGCGTGGCT 1521
DB 479 ProLysGluSerGlnArgValValAla 488

RESULT 11
US-10-425-114-44047
; Sequence 44047, Application US/10425114
; Publication NO. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
```

; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 44047
 ; LENGTH: 491
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700763253_FLI.pap
 ; US-10-425-114-44047

Alignment Scores:
 Pred. No.: 2,01e-120 Length: 491
 Score: 1502.00 Matches: 297
 Percent Similarity: 71.57% Conservatism: 68
 Best Local Similarity: 58.24% Mismatches: 91
 Query Match: 55.82% Indels: 54
 DB: 15 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-425-114-44047 (1-491)

QY 1 CGGAGCAACCGGAGGAGAGCGGTGCGAGTTCAGCGCCTCAATGCGCAGAGACCTGAC 60
 DB 33 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 52
 QY 61 AATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTTGGAACCCCAACACAGGAGTTC 120
 DB 53 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 72
 QY 121 GAATTCGCCCGCGGTGCGCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCCT 180
 DB 73 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 92
 QY 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCAGCAAGAGGAGGATCTTTGGGTTG 240
 DB 93 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 112
 QY 241 ATATTCCCTGTTGCTTACAGACTATGAGAGCCTCAGACACAGAGTCTGCTATCTCAG 300
 DB 113 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 129
 QY 301 TCCCAAGACACCAAGACGCTCTCCAGGAGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
 DB 130 SerSerArgPro-----GlnAspArg 136
 QY 361 CACAGAGGTTCACCGGTTTCGATGAGGGTGAATCTCATTCAGTTCACCGGTGTTGCT 420
 DB 137 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 156
 QY 421 TTCTGGCTCTACACGACGACGACACTGATGTTGTTGCTGTTCTTCTTACTGACACCAAC 480
 DB 157 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 176
 QY 481 AACACGACACCAAGCTTGTATGATCTCCAGGAGATTCATATTTGGTGGGAAACAGGAG 540
 DB 177 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 196
 QY 541 CAGAGGTTCTTAGGTACAGCAACCAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
 DB 197 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 209
 QY 601 AGCCCATACAGCCCGCAAGTCTGAGCTAGACAGAGAGCGTGNATTTAGCCCTCAGGA 660
 DB 210 -----Gly 210
 QY 661 CAGCAGACCGCAGAGAAACGAGCAGGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 DB 211 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 230
 QY 721 AGCGGTTCTACCGCGGAGTTCTTGGAAACAAAGCTTCCAGGTTCCAGCAGACAGATAGTG 780
 DB 231 SerGlyPheAlaProGluPheLeuGluHieAlaPheValVal---AspArgGlnIleVal 249
 QY 781 CAAAACCTAAGAGCGGAGCCGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

Db 250 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyAlaIleValThrValLysGly 269
 QY 841 GGCTCTCAGAAATCTTGAAGCCCA-----GATGAAAGAGAGCTGCCAGCAAGAGAG 891
 Db 270 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 289
 QY 892 GAATACGATGAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 Db 290 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 304
 QY 952 AGAGGCGAGGGGGAATGTTGATGAGAGAGAGATGTCACCGCAAGTCTGTAAGAGACATT 1011
 Db 305 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 321
 QY 1012 GGTAGAAACAGATCCCTCGATCTACAACTCAAGCTTCAAGCTTCACTCAAACTGCCAAC 1071
 Db 322 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 341
 QY 1072 GATCTCAACCTTCTAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1131
 Db 342 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 361
 QY 1132 AGGAATGCTATTGTTCTGCTCACTACACACCAAGCAGCAGCAGCATCATATATCGATTG 1191
 Db 362 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 381
 QY 1192 AGGGGAGCGGCTCACGTGCAAGTCTGAGCAGCAACGCAACAGAGTGTACGACGAGGAG 1251
 Db 382 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 401
 QY 1252 CTTCAAGAGGTCACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311
 Db 402 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 421
 QY 1312 AGCGAAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1371
 Db 422 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 441
 QY 1372 GCGGTGTGAAATCTCCCTCATAGTACCTGCGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCT 1431
 Db 442 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnThrPheAsn 461
 QY 1432 CTTCAAGAGGAGCAGCAGGAGCTTAAAGAACAAACACCCCTTCAAGTCTTCTGTTCCA 1491
 Db 462 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 481
 QY 1492 CGGTCTCAGCAGTCTCCGAGGCGCTGCTGCT 1521
 Db 482 ProLysGluSerGlnArgValValAla 491

RESULT 12

US-10-425-114-43836
 ; Sequence 43836, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425, 114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 43836
 ; LENGTH: 492
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:

OTHER INFORMATION: Clone ID: 700673875_FLI.pap
US-10-425-114-43836

Alignment Scores: 2,01e-120 Length: 492
Pred. No.: 1502.00 Matches: 297
Score: 71.57% Conservative: 68
Percent Similarity: 58.24% Mismatches: 91
Best Local Similarity: 55.82% Indels: 54
Query Match: 15 Gaps: 7
DB:

US-10-728-323-3 (1-1524) x US-10-425-114-43836 (1-492)

```

QY 1 CGGACGACCGGAGAGACGGTCCAGGCTTCCAGCGCTCAATCGCGAGACCTGAC 60
DB 1 CGGACGACCGGAGAGACGGTCCAGGCTTCCAGCGCTCAATCGCGAGACCTGAC 60
QY 34 ArgGluGlnProGlnGlnAenGluCysGlnGlnArgLeuAenAlaLeuLysProAep 53
DB 34 ArgGluGlnProGlnGlnAenGluCysGlnGlnArgLeuAenAlaLeuLysProAep 53
QY 61 AATCGATTGAATCAGAGCGCGTTCATGAGACTTGGACCCCAACACGAGAGTTC 120
DB 61 AATCGATTGAATCAGAGCGCGTTCATGAGACTTGGACCCCAACACGAGAGTTC 120
QY 54 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAenAenLysProPhe 73
DB 54 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAenAenLysProPhe 73
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCTAGGCT 180
DB 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCTAGGCT 180
QY 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAenArgAenAlaLeuArgPro 93
DB 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAenArgAenAlaLeuArgPro 93
QY 181 TTCTACTTCCAAATGCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGAGTACTTTGGGTTG 240
DB 181 TTCTACTTCCAAATGCTCCCGAGGAGATCTTTCATCCAGCAAGGAGGAGTACTTTGGGTTG 240
QY 94 SerTyThrAenAlaProGlnGluIleTyIleGlnGlnGlySerGlyIlePheGlyMet 113
DB 94 SerTyThrAenAlaProGlnGluIleTyIleGlnGlnGlySerGlyIlePheGlyMet 113
QY 241 ATATTCCTGGTGTCTAGACACTATGAAGAGCTTACACACAGAGTGTGTGATCTCAG 300
DB 241 ATATTCCTGGTGTCTAGACACTATGAAGAGCTTACACACAGAGTGTGTGATCTCAG 300
QY 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlyGlyGlyGlyGly 130
DB 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlyGlyGlyGlyGly 130
QY 301 TCCCAAGACACCAAGAGCTCTCCAGGAGAGAGACCAAGCCCAAGCAAGAGATAGT 360
DB 301 TCCCAAGACACCAAGAGCTCTCCAGGAGAGAGACCAAGCCCAAGCAAGAGATAGT 360
QY 131 SerSerArgPro-----GlnAspArg 137
DB 131 SerSerArgPro-----GlnAspArg 137
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
DB 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
QY 138 HisGlnLysIleTyHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
DB 138 HisGlnLysIleTyHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
QY 421 TTCTGGCTTACACGACACGACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480
DB 421 TTCTGGCTTACACGACACGACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480
QY 158 TyrTrpMetTyThrAenAenGluAspThrProValAlaValSerLeuIleAspThrAen 177
DB 158 TyrTrpMetTyThrAenAenGluAspThrProValAlaValSerLeuIleAspThrAen 177
QY 481 AACACGACACCAAGCTTGATCATGTTCCCGAGGAGATTCATTTGCTGGGACACGGAG 540
DB 481 AACACGACACCAAGCTTGATCATGTTCCCGAGGAGATTCATTTGCTGGGACACGGAG 540
QY 178 SerPheGlnAenGlnLeuAenGlnMetProArgArgPheTyThrLeuAlaGlyAenGlnGlu 197
DB 178 SerPheGlnAenGlnLeuAenGlnMetProArgArgPheTyThrLeuAlaGlyAenGlnGlu 197
QY 541 CAAGAGTTCTTAAGGTACACGACCAACAGCAGACGACGACGACGACGACGACGACGAC 600
DB 541 CAAGAGTTCTTAAGGTACACGACCAACAGCAGACGACGACGACGACGACGACGACGAC 600
QY 198 GlnGluPheLeuGlnTyThrGlnProGlnLysGlnGlnGly-----Gly 211
DB 198 GlnGluPheLeuGlnTyThrGlnProGlnLysGlnGlnGly-----Gly 211
QY 661 CAGCAGACGCGCAGAGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720
DB 661 CAGCAGACGCGCAGAGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720
QY 212 ThrGlnSerGlnLysGlyTyArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 231
DB 212 ThrGlnSerGlnLysGlyTyArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 231
QY 721 AGCGGCTTACGCGCGGAGTTCTTCGGAACCAAGCCCTTCAGGTTTCAGCAGCAGACATAGTG 780
DB 721 AGCGGCTTACGCGCGGAGTTCTTCGGAACCAAGCCCTTCAGGTTTCAGCAGCAGACATAGTG 780
QY 232 SerGlyPheAlaProGluPheLeuGluHieAlaPheValVal---AspArgGlnIleVal 250
DB 232 SerGlyPheAlaProGluPheLeuGluHieAlaPheValVal---AspArgGlnIleVal 250
QY 781 CAAAACCTAAGAGCGGACCGCAGAGTAGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 CAAAACCTAAGAGCGGACCGCAGAGTAGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 251 ArgLysLeuGlnGlyGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 270
DB 251 ArgLysLeuGlnGlyGluAenGluGluGluGluGluGluGluGluGluGluGluGluGlu 270
QY 841 GGCTCAGAACTTTGAGCCCA-----GATGAAAGAGAGCTGCCGACGAGAGAGAG 891
DB 841 GGCTCAGAACTTTGAGCCCA-----GATGAAAGAGAGCTGCCGACGAGAGAGAG 891
QY 271 GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 290
DB 271 GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 290

```

Alignment Scores: 2,01e-120 Length: 492
Pred. No.: 1502.00 Matches: 297
Score: 71.57% Conservative: 68
Percent Similarity:

```

QY 892 GAATACGATGAAGATGAATATGATATCGATGAGAGGATAGAGCGGTGGCAGGGAAGC 951
DB 892 GAATACGATGAAGATGAATATGATATCGATGAGAGGATAGAGCGGTGGCAGGGAAGC 951
QY 291 LysProAep-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
DB 291 LysProAep-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
QY 952 AGAGCAGCGGGGAATGTTATTGAAGAGAGCATCTCCACCGCAAGTCTTAAAGAAAGAACATT 1011
DB 952 AGAGCAGCGGGGAATGTTATTGAAGAGAGCATCTCCACCGCAAGTCTTAAAGAAAGAACATT 1011
QY 306 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 322
DB 306 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 322
QY 1012 GGTAGAAACAGATCCCTGACATCTACCAACCTCAAGCTGTGTCTCACTCAAACTCCCAAC 1071
DB 1012 GGTAGAAACAGATCCCTGACATCTACCAACCTCAAGCTGTGTCTCACTCAAACTCCCAAC 1071
QY 323 GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrAlaThr 342
DB 323 GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrAlaThr 342
QY 1072 GATCTCAACCTTCTAATTAAGTGGCTTGGACCTAGTCTGCTGAATATGGAATCTCTAC 1131
DB 1072 GATCTCAACCTTCTAATTAAGTGGCTTGGACCTAGTCTGCTGAATATGGAATCTCTAC 1131
QY 343 SerLeuAaspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362
DB 343 SerLeuAaspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362
QY 1132 AGGAATGATCTGTTGTCCTCCTACCAACCAACGACACAGCATCATATATCGATTG 1191
DB 1132 AGGAATGATCTGTTGTCCTCCTACCAACCAACGACACAGCATCATATATCGATTG 1191
QY 363 LysAenAlaMetPheValProHisTyThrAenLeuAenAlaAenSerIleIleTyAlaLeu 382
DB 363 LysAenAlaMetPheValProHisTyThrAenLeuAenAlaAenSerIleIleTyAlaLeu 382
QY 1192 AGGGACCGGCTTCAGTCCAGTCTGACAGTCTGACAGCAACGCGCAACAGAGTGTCACGAGGAG 1251
DB 1192 AGGGACCGGCTTCAGTCCAGTCTGACAGTCTGACAGCAACGCGCAACAGAGTGTCACGAGGAG 1251
QY 383 AsnGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGlyGlu 402
DB 383 AsnGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGlyGlu 402
QY 1252 CTTCAAGAGGTCACGCTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAGTCCAG 1311
DB 1252 CTTCAAGAGGTCACGCTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAGTCCAG 1311
QY 403 LeuGlnGluGlyGlnValLeuIleValProGlnAenPheAlaValAlaAalargSerGln 422
DB 403 LeuGlnGluGlyGlnValLeuIleValProGlnAenPheAlaValAlaAalargSerGln 422
QY 1312 ACCGAGAACTTCGAATAGTGCATTCAAGCAGACACTCAAGGCCAGCATAGCAACCTC 1371
DB 1312 ACCGAGAACTTCGAATAGTGCATTCAAGCAGACACTCAAGGCCAGCATAGCAACCTC 1371
QY 423 SerAspAenPheGluTyThrValSerPheTyThrAenAaspArgProSerIleGlyAenLeu 442
DB 423 SerAspAenPheGluTyThrValSerPheTyThrAenAaspArgProSerIleGlyAenLeu 442
QY 1372 GCCGTGTGAAACTCCGTCTATAGATAACCTCCGAGGAGGCTGTTGCCAAATTCATATGTC 1431
DB 1372 GCCGTGTGAAACTCCGTCTATAGATAACCTCCGAGGAGGCTGTTGCCAAATTCATATGTC 1431
QY 443 AlaGlyAlaAenSerLeuLeuAenAlaLeuProGluGluValIleGlnGlnThrPheAen 462
DB 443 AlaGlyAlaAenSerLeuLeuAenAlaLeuProGluGluValIleGlnGlnThrPheAen 462
QY 1432 CTTCAAGAGGACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491
DB 1432 CTTCAAGAGGACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491
QY 463 LeuArgArgGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro 482
DB 463 LeuArgArgGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro 482
QY 1492 CGGTCTCAGCATCTCCGAGGCGCTGGCT 1521
DB 1492 CGGTCTCAGCATCTCCGAGGCGCTGGCT 1521
QY 483 ProLysGluSerGlnArgArgValValAla 492
DB 483 ProLysGluSerGlnArgArgValValAla 492

```

RESULT 13
US-10-425-114-46294
Sequence 46294, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 46294
LENGTH: 492
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700757865_FLI.pap
US-10-425-114-46294

Best Local Similarity: 58.24%		MisMatches: 91	
Query Match: 55.82%		Indels: 54	
DB: 15		Gaps: 7	
US-10-728-323-3 (1-1524) x US-10-425-114-46294 (1-492)			
QY	1	CGGAGCAACCGGAGGAGAACCGGTGCCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
DB	34	ArgGluGlnProGlnGlnAenGluCysGlnIleGlnArgLeuAenAlaLeuLysProAsp	53
QY	61	AATCGCATTAATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACGAGGATTC	120
DB	54	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAenProAenAenLysProPhe	73
QY	121	GAATGCGCGCGGTCCGCTCTCTCGCTTAGTCTCGCGCCCAACGCGCTTCGTAGGCCT	180
DB	74	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAenArgAenAlaLeuArgPro	93
QY	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAAGGGATATCTTTGGGTG	240
DB	94	SerTyrThrAenAlaProGlnGluIleTyrIleGlnGlySerGlyIlePheGlyMet	113
QY	241	ATATTCCCTGGTGTCTAGACACTATGAAGAGCTCTACACAAAGTCTCGTGCATCTCAG	300
DB	114	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	130
QY	301	TCCCAAGACACCAAGACGCTCTCCAAGGAGAGAACCAAGCCAAACAGCAACGAGATAGT	360
DB	131	SerSerArgPro-----GlnAspArg	137
QY	361	CACCAAGGTGCACCGTTTCGATGAGGGTATCTCATTCAGTCTCCACCGGTGTGCT	420
DB	138	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	157
QY	421	TTCTGGCTCTAACGACACGACACTGATGTGTGTCTCTTCTTACTGACACCAAC	480
DB	158	TyrTrpMetTyrAenAenGluAepThrProValValAlaValSerLeuIleAepThrAen	177
QY	481	AACAAACGACACGAGTGTGATCAGTTCGCCAGGAGATTCAATTTGGTGGGAACACGGAG	540
DB	178	SerPheGlnAenGlnLeuAepGlnMetProArgArgPheTyrLeuAlaGlyAenGlnGlu	197
QY	541	CAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT	600
DB	198	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly-----Gly	210
QY	601	AGCCCATACAGCCCGCAAGTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	211	-----Gly	211
QY	661	CAGCAGACGCGCAGAGAACGAGCAGGACGAGCAAGAAAGAAACGAAAGTGAACATCTTC	720
DB	212	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluAenGluGlyGlySerIleLeu	231
QY	721	AGCGGTTCCCGCGGAGTCTCTGGAACAAAGCTTCCAGGTTGACGACAGACGATAGTG	780
DB	232	SerGlyPheAlaProGluPheLeuGluHisAlaPheVal---AspArgGlnIleVal	250
QY	781	CAAAACCTTAAGAGGGAGACCGAGAGTGAAGAGAGGGAGCCATTGTGACAGTGAGGGA	840
DB	251	ArgLysLeuGlnGlyGluAenGluGluGluGlyAlaIleValThrValLysGly	270
QY	841	GGCTCAGAACTCTGAGCCCA-----GATAGAAAGACGTCGCCGACGAAGAAGAG	891
DB	271	GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGluGlu	290
QY	892	GAATACGATGAAGATGAATATGATACGATGAAGAGGATAGAGCGGTGGCAGGGAAGC	951
DB	291	LysProAep-----CysAspGluLysAspLysHisCysGlnSerGlnSer	305
QY	952	AGAGCGAGGGGAATGTTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACATT	1011
DB	306	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	322
QY	1012	GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGGTTCACCTCAAAACGCCAAC	1071
DB	323	GlyGlnThrSerSerProAspIlePheAenProGlnAlaGlySerIleThrThrAlaThr	342
QY	1072	GATCTCAACTCTTAATACTAGTGGCTTGGACCTAGTGTGAATATGGAATCTCTAC	1131
DB	343	SerLeuAepPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	362
QY	1132	AGAAATGCAATGTTTCTCGCTCACTACAACCAACGACACAGCATCATATATCGATTG	1191
DB	363	LysAenAlaMetPheValProHisTyrAenLeuAenAlaAenSerIleIleTyrAlaLeu	382
QY	1192	AGGGAGCGGCTCACGTGCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAG	1251
DB	383	AsnGlyArgAlaLeuValGlnValValAenCysAenGlyGluArgValPheAspGlyGlu	402
QY	1252	CTTCAAGAGGTCACGTCTGTGGTCCACAGAACTTCGCGCTCGTGGAAAGTCCCAG	1311
DB	403	LeuGlnGluGlyGlnValLeuIleValProGlnAenPheAlaValAlaAenSerGln	422
QY	1312	AGCGAAGACTTCGAATACGTGGCTTCAAGACAGACTCAAGGCCAGCATAGCAACCTC	1371
DB	423	SerAspAenPheGluTyrValSerPheLysThrAenAepArgProSerIleGlyAenLeu	442
QY	1372	GCGGTGAAAACTCCGTCTAGATAAACCCTGCGGAGGAGGTGGTTGCCAATTCATATGGC	1431
DB	443	AlaGlyAlaAenSerLeuAenAlaLeuProGluGluValIleGlnGlnThrPheAen	462
QY	1432	CTCCAAGGAGGAGCAGGCAAGCGAGCTTAAAGAACCAACCCCTTCAAGTCTTCGTTC	1491
DB	463	LeuArgArgGlnGlnAlaArgGlnValLysAenAenAenProPheSerPheLeuValPro	482
QY	1492	CCGTCTCAGCAGTCTCCGAGGCGCTGTGGCT	1521
DB	483	ProLysGluSerGlnArgArgValValAla	492
RESULT 14			
US-10-425-114-71959			
; Sequence 71959, Application US/10425114			
; Publication No. US2004003488A1			
; GENERAL INFORMATION:			
; APPLICANT: Liu, Jingdong			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Screen, Steven E			
; APPLICANT: Tabaska, Jack E			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With			
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement			
; FILE REFERENCE: 38-21(5313)B			
; CURRENT APPLICATION NUMBER: US/10/425,114			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 73128			
; SEQ ID NO 71959			
; LENGTH: 466			
; TYPE: PRT			
; ORGANISM: Glycine max			
; FEATURE:			
; OTHER INFORMATION: Clone ID: 700676992_FLI.pep			
US-10-425-114-71959			
Alignment Scores:			
Pred. No.:		3,23e-120	Length: 466
Score:		1499.50	Matches: 297
Percent Similarity:		72.02%	Conservative: 71
Best Local Similarity:		58.12%	Mismatches: 94
Query Match:		55.72%	Indels: 49
DB:		15	Gaps: 6
US-10-728-323-3 (1-1524) x US-10-425-114-71959 (1-466)			
QY	1	CGGAGCAACCGGAGGAGAACCGGTTCAGCGCTCAATGCGCAGACCTGAC	60


```
Db 1 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
  |||::||| :::::|||| ||||| |||::||| ||||| |||::|||
Qy 61 AATCGCATTAATCAGAGCGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
  ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 GAATGGCGCGGCTCGCGCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
  ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 60
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 TTCTACTCCAATGCTCCCAAGAGATCTTCATCCAGCAAGGAAGGATCTTTGGGTG 240
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 80
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 241 ATATTCCCTGTTGCTAGACACTATGAAGAGCCTCACACAAAGTCTGCTCATCTCAG 300
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 100
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 301 TCCCAAGACCAACCAAGACGCTCCCAAGGAGAGACCAAGCCCAACAGCAACGAGATAGT 360
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 101 SerGlnArgPro-----GlnAspArg 107
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 361 CACAGAAGTGTACCGTTTCGATGAGGGTGATCTCATTCAGTTCACCGCGGTGTGCT 420
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 421 TTCTGGCTTACACGACCAACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 147
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 481 AACAAACACACACGAGTTCAGTTCCTCCCGAGAGATTCAAATTTGGCTGGGAACGAG 540
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 148 SerLeuGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 167
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 541 CAAGAGTCTTAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTACCATAT 600
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 168 GlnGluPheLeuLysTyrGlnGlnGln----- 176
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 601 AGCCCATACGCCCGCAAGCTAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGA 660
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 177 -----GlnGlnGlyGlySerGlnSerGlnLysGly 186
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 661 CAGCAGACCGCGCAGAGAACGAGCAGACGAAGAAAGAAAGAGTGGAAACATCTTC 720
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 LysGln-----GlnGluGluAsnGluGlySerAsnIleLeu 199
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 AGCGGTTCACGCCGAGTTCCTGGAAACAGCCTTCAGGTTGACACAGACAGATAGTG 780
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 218
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 CAAACCTAAGAGCGGACCGAGAGTGAGAGAGGAGGCCATTCGACAGTGGAGGGA 840
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 ArgAsnLeuGlnGlyGluAsnGluGluGluAspSerGlyAlaIleValThrValLysGly 238
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 GGCTCAGAACTTTGAGCCAGATAGAAGAGACGTCGCCAGCAAGAGAGAAATACGAT 900
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 901 GAAGATGAA-----TATGAATACGATGAAGAGATAGAAGCGTGGCAGGGGA 948
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 949 AGCAGAGGCGGGGATGTTTGAAGAGACCATCTGCACCCCAAGTGTCTAAAGAAC 1008
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1009 ATTTGGTAGAAACAGATCCCTGACATCTACAACCCCTCAAGCTGTTTCACTCAAAACTGCC 1068
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 296 IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 315
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1069 AACGATCTCAACCTTCTAATACCTTAGTGCGCTGGACCTAGTCTGAAATATGAATCTC 1128
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1129 TACAGGAATGCATTGTTTGTCTCACTACACACCAACGACGACATCATATATCGA 1188
  ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 355
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1189 TTGAGGGGACGGGCTCAGCTGCAAGTCTGGAGCAACCGCAACAGAGTGTACGACGAG 1248
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1249 GAGCTTCAAGAGGTCAGCTGCTTGTGTGTCACAGAACTTCGCGTGTGCGAAGTCC 1308
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 376 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 395
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1309 CAGAGCGGAACCTCGAATACGTGCATTCAAGACAGACTCAAGGCCAGCATAGCCAAC 1368
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 396 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1369 CTCGCGGTGAAAACTCCGTATAGATAACCTGCCGAGGAGGTGGTTGCATAATTCATAT 1428
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 LeuAlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 435
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1429 GGCCTCCAAAGGAGCAGCGCAGGCTTAAGAACAAACCCCTTCAAGTCTTCGTT 1488
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 455
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1489 CCACCGTCTCAGCAGTCTCCGAGGCGCTGTGGCT 1521
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 456 ProProGlnGluSerGlnArgArgAlaValAla 466
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 15
US-10-425-114-71973
; Sequence 71973, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71973
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700677092_FLI.pep
US-10-425-114-71973
```

```
Alignment Scores:
Pred. No.: 3-23e-120 Length: 466
Score: 1499.50 Matches: 297
Percent Similarity: 72.02% Conservative: 71
Best Local Similarity: 58.12% Mismatches: 94
Query Match: 55.72% Indels: 49
DB: 15 Gaps: 6
```

US-10-728-323-3 (1-1524) x US-10-425-114-71973 (1-466)

```
Qy 1 CCGCAGACACCGAGAGAGAACCGCTGCCAGTTCACGCGCTCAATCCGACGAGACCTGAC 60
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 20
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 AATCCGATTAATCAGAGCGGTTACATTGAGACTTGGAAACCCCAACACGAGGATTC 120
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 AsnA-gileGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 40
  |||::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


This Page Blank (uspto)